

# OPTIMUM AND EQUILIBRIUM IN A TRANSPORT PROBLEM WITH QUEUE PENALIZATION EFFECT

GIANLUCA CRIPPA, CHLOÉ JIMENEZ, AND ALDO PRATELLI

**ABSTRACT.** Consider a distribution of citizens in an urban area in which some services (supermarkets, post offices. . . ) are present. Each citizen, in order to use a service, spends an amount of time which is due both to the travel time to the service and to the queue time waiting in the service. The choice of the service to be used is made by every citizen in order to be served more quickly. Two types of problems can be considered: a global optimization of the total time spent by the citizens of the whole city (we define a global *optimum* and we study it with techniques from optimal mass transportation) and an individual optimization, in which each citizen chooses the service trying to minimize just his own time expense (we define the concept of *equilibrium* and we study it with techniques from game theory). In this framework we are also able to exhibit two time-dependent strategies (based on the notions of prudence and memory respectively) which converge to the equilibrium.

## 1. INTRODUCTION

In this paper we study an optimization problem coming from the modeling of the behaviour of citizens who every day need to use some services (supermarkets, post offices. . . ) present in a city. We consider a bounded set  $\Omega \subset \mathbb{R}^d$  which is the geographic reference for the city and a nonnegative function  $f$  with unit integral representing the population density. We fix  $k$  points  $x_1, x_2, \dots, x_k$  at which the services are located. If every service was able to answer immediately to the demand of each customer, obviously every person would choose the one closest to his home. Namely, being  $p \geq 1$  a fixed number so that the time spent to cover a distance  $\ell$  is given by  $\ell^p$ , then a citizen living at  $x$  would choose the service located at  $x_i$  if and only if

$$|x - x_i|^p = \min_{j=1, \dots, k} |x - x_j|^p.$$

However, in real world, if a service is crowded because a certain amount of citizens have chosen it, the satisfaction of the demand of the customer is not immediate, and some time has to be spent waiting for it. This time spent in the queue surely depends on the amount of people waiting at the service, but also on the characteristics of the service (for instance, its dimension or the ability of the employees). This can be modeled using  $k$  functions  $h_1, h_2, \dots, h_k$  which express the time to be waited in dependence of the amount of people in the queue. If the amount of citizens choosing the service  $x_j$  is  $c_j$ , for  $j = 1, \dots, k$ , then a citizen living at  $x$  and going to the service  $x_i$  needs to use a total amount of time expressed by the quantity

$$|x - x_i|^p + h_i(c_i)$$

in order to reach the service  $x_i$  and to have his request fulfilled. In particular, the better service for the customer is not necessarily the closest one. It could be convenient for him to go a bit further away in order to be served in a bigger service, in which the queue is faster.

It is surely apparent from the above discussion that, in this modeling, the decision of each citizen depends on the choices of all the other citizens. Each customer chooses the best service

according to the corresponding queue, but the queue itself depends on the choices of all the other customers. A possible attempt to tackle this question is to look for a global minimum. Let us denote by  $A_j$  the subset of  $\Omega$  consisting of the areas in which the citizens using the service  $x_j$  live. The total cost, in term of time spent to be served, needed by the citizens of the whole city is expressed by

$$\sum_{i=1}^k \int_{A_i} \left[ |x - x_i|^p + h_i \left( \int_{A_i} f(x) dx \right) \right] f(x) dx .$$

If we agree that the choice of each citizen is done in order to minimize this total cost (for instance, if the major of the city has the right to force the choices of the customers, in order to make the global cost as small as possible), we are led to consider the following minimization problem:

$$\inf_{\substack{(A_i)_{i=1,\dots,k} \\ \text{partition of } \Omega}} \left\{ \sum_{i=1}^k \int_{A_i} \left[ |x - x_i|^p + h_i \left( \int_{A_i} f(x) dx \right) \right] f(x) dx \right\} . \quad (1.1)$$

This type of minimization problem fits very well in the theory of optimal mass transportation. This theory goes back to Monge [25], who considered the problem of moving a pile of sand into a hole of the same volume, minimizing the transportation cost. We briefly refer on the Monge problem, its relaxation due to Kantorovich ([21] and [22]) and the recent progresses on it in Section 2.2. The study of the minimization problem (1.1) is done in Section 3. We define the concept of *optimum*, which is a partition  $(A_i)_{i=1,\dots,k}$  solving (1.1), we prove results of existence and uniqueness of the optimum (under suitable assumptions on the functions  $h_i$ ) and we give some characterizations of it.

However, this kind of solution of the problem is not very natural, from an economic point of view. Indeed, as pointed out in Example 4.7, it could happen that for some citizens the choice of the service forced by the need of producing an optimum for (1.1) is too expensive. It is more realistic to consider as a good solution of the problem a partition of  $\Omega$  in which each citizen is “satisfied” with his own choice. By this we mean the following: the citizen looks at the behaviour of all the other citizens, and regarding it as fixed he decides whether his decision has been clever or not. In particular, a citizen living at  $x$  is satisfied with his choice of going to the service located at  $x_i$  if and only if

$$|x - x_i|^p + h_i(c_i) = \min_{j=1,\dots,k} \{ |x - x_j|^p + h_j(c_j) \} , \quad (1.2)$$

where the numbers  $c_j$ , as before, represent the amount of customers choosing the service  $x_j$ , for  $j = 1, \dots, k$ , but now these quantities are seen as fixed by the citizen who evaluates the correctness of his decision. This is radically different from the previous minimization problem, as we explicitly remark in Section 4.4. The condition expressed in (1.2) paves the way for a connection with game theory (for which good references are [18] and [5]): in Section 4 we define an *equilibrium* for our problem as a partition in which every customer is satisfied with his choice, in the sense specified above. In classical game theory this corresponds to the so-called Nash equilibrium. We are able to prove, under suitable assumptions on the functions  $h_i$ , existence and uniqueness of an equilibrium, and we also comment on the main differences between equilibrium and optimum.

In Section 5 we finally consider a dynamical evolution of the problem, looking for convergence to the equilibrium. The general question is the following. Assume that the citizens have some strategy to decide, day by day, the most convenient service, using data which come from the

previous days. At each day, they decide where to go, considering the lengths of the queues that have been observed in the previous days. Does this iterative scheme converge to the equilibrium? In Example 5.3 we show that this is not the case, if the strategy of the customers is too naïve. If they just choose every day to minimize the right hand side of (1.2), in which the time spent on the queues in the various services is simply the one observed the day before, then we could obtain some oscillatory phenomena, with no convergence. However, we propose two different strategies, quite natural at a social level, which lead to convergence to the equilibrium.

The first one (see Section 5.2) is based on the concept of *prudence* (which, in some sense, could also be viewed as laziness or resistance to a change of habits). The decision that every citizen makes is again based on the information on the queues of the previous day, but it is now seen as a stochastic variable (in the language of optimal mass transportation this corresponds to the notion of transport plan, see again Section 2.2). It is “too risky” for the customer to change immediately the preferred service, thus we assume that he is only going to change his opinion, in the sense that he will be more tempted by the most convenient service, but he will not immediately “fully choose” it. The choice between two services of each customer is modeled by a function with values between 0 and 1: value 0 means that the choice is completely in favour of the first service, while value 1 means that the choice is completely in favour of the second service. Values between 0 and 1 mean that the customer chooses a mixed strategy (this is also common in game theory). This can also be interpreted in statistical terms: if at each point of the city there is a condominium instead of a single citizen, then each person will choose the service in such a way that the statistical distribution of the choices follows the stochastic variable defined above. Assuming that the citizens have “enough prudence”, we are able to show that the iteration of the stochastic choices converges to a deterministic decision, which is an equilibrium for our problem.

The second strategy (see Section 5.3) relies on the notion of *memory*. We go back to a deterministic strategy, but we assume that each customer bases his decision not only on the informations coming from the previous day, but also remembering the queues that have been seen in a certain amount of days in the past. Then the choice for the new day is based on an average of all these informations. If the citizens have “enough memory”, in the sense that they consider in their choice a sufficiently big amount of days, we are able to show that also this iteration converges to an equilibrium for the problem.

We close this introduction by relating our problem to some other optimization problems studied in the previous literature. The idea for our model came from some results relative to the so-called *location problem*, see [6], [20] and [7]. In this problem, we are given a distribution of citizens  $f$  in a city  $\Omega$ , but the location of the services is no more fixed. The question is to optimize the location of the services, in such a way that the transportation cost of the citizens onto the services is minimized. We observe that no queues are considered in this model. Two different approaches to this problem are possible. In the *long term planning* all the services are built at the same time, minimizing the average distance the people have to cover to reach the nearest service. In the *short term planning*, it is assumed that services are built one by one, minimizing the average distance step by step, taking into account at each stage the location of all the services opened at previous steps. Other related works, regarding the modeling of urban areas and the structure of cities, are [9], [11], [14], [15], [13] and [24].

**Acknowledgment.** The authors warmly thank Luigi Ambrosio, Guy Bouchitté, Guillaume Carlier, Fabio Priuli and Filippo Santambrogio for many valuable conversations on the subject of

this work. Moreover, they also thank the Scuola Normale Superiore di Pisa and the Mathematics Department of the University of Pavia for the hospitality during the preparation of this paper.

## 2. NOTATION AND PRELIMINARY RESULTS ON OPTIMAL MASS TRANSPORTATION

In this section we introduce the general setup of the paper and the main notation, we review some basic facts about optimal mass transportation (for which the reader can consult [1] or [27]) and we establish some preliminary results which will be useful in the sequel.

**2.1. Main notation of the paper.** We fix a bounded Borel set  $\Omega \subset \mathbb{R}^d$  and we denote by  $\mathbb{L}^d$  the  $d$ -dimensional Lebesgue measure in  $\mathbb{R}^d$ . We consider  $k$  points  $x_1, x_2, \dots, x_k$  belonging to  $\Omega$ , where  $k$  is a strictly positive integer; these points represent the location of the services in the city. The density of the citizens is given by an absolutely continuous probability measure  $\mu = f\mathbb{L}^d \llcorner \Omega$ , where  $f : \Omega \rightarrow \mathbb{R}$  is a nonnegative function with unit integral. The measure  $\mu$  defined in this way will be typically the reference measure in  $\Omega$ , and we shall also write  $f$ -a.e. and  $f$ -negligible as a shortening of the more precise notation  $\mu$ -a.e. and  $\mu$ -negligible. For  $i = 1, \dots, k$  we consider the functions  $h_i : [0, 1] \rightarrow [0, +\infty[$  which encode the amount of time to be waited in dependence of the amount of people using the service. With this we mean that, if the service located at  $x_i$  is chosen by an amount  $c_i$  of citizens, then the time that will be spent in the queue by each customer is  $h_i(c_i)$ . Notice that in this modeling we can also include some penalizations for the particular features of the various services, choosing the queue functions in such a way that  $h_i(0) > 0$ : for instance, higher prices or lower quality of the products. We do not specify a priori any condition on the functions  $h_i$ , but we will rather clarify the assumptions needed in each particular result. For every Borel set  $A \subset \Omega$  we consider the indicatrix function  $\mathbf{1}_A$  defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

The set of probability measures on  $\Omega$  and  $\Omega \times \Omega$  are denoted by  $\mathcal{P}(\Omega)$  and  $\mathcal{P}(\Omega \times \Omega)$  respectively. When  $\gamma \in \mathcal{P}(\Omega)$  is given, we denote by  $L_\gamma^1(\Omega)$  the vector space consisting of the Borel functions which are  $\gamma$ -integrable. A partition of  $\Omega$  is a finite or countable family  $(A_i)_{i \in \mathbb{N}}$  of pairwise disjoint (up to  $f$ -negligible sets) Borel sets  $A_i \subset \Omega$  such that  $\cup_{i \in \mathbb{N}} A_i$  has full measure in  $\Omega$ .

**2.2. Optimal mass transportation.** In 1781 Monge [25] raised the problem of transporting a given distribution of mass (a pile of sand, for instance) into another one (a hole, for instance), in such a way that the total work done is minimal. In modern terms the problem can be stated as follows. Given two measures  $\mu, \nu \in \mathcal{P}(\Omega)$  and  $p \geq 1$  we consider the minimization problem

$$M_p(\mu, \nu) = \inf \left\{ \left( \int_{\Omega} |x - T(x)|^p d\mu(x) \right)^{1/p} : T : \Omega \rightarrow \Omega \text{ Borel and such that } T_{\#}\mu = \nu \right\}.$$

Here we denote by  $T_{\#}\mu \in \mathcal{P}(\Omega)$  the push-forward of the measure  $\mu$ , defined by

$$(T_{\#}\mu)(A) = \mu(T^{-1}(A)) \quad \text{for every Borel set } A \subset \Omega.$$

Any map  $T$  which is admissible in the above problem is called a transport map from  $\mu$  to  $\nu$ .

It is extremely difficult to attack the Monge problem, mainly due to the fact that it is highly nonlinear. This is why the following relaxed formulation, due to Kantorovich ([21] and [22]), is of great importance. We denote by  $\Pi(\mu, \nu) \subset \mathcal{P}(\Omega \times \Omega)$  the set of the probability measures  $\gamma$  in  $\Omega \times \Omega$  with marginals  $\mu$  and  $\nu$ , i.e. such that  $(\pi_1)_{\#}\gamma = \mu$  and  $(\pi_2)_{\#}\gamma = \nu$ , where we denote by  $\pi_1$  the projection on the first component and by  $\pi_2$  the projection on the second component.

Every element of  $\Pi(\mu, \nu)$  is called a transport plan from  $\mu$  to  $\nu$ . We notice that  $\Pi(\mu, \nu)$  is always nonempty, since for instance  $\mu \otimes \nu \in \Pi(\mu, \nu)$ . The relaxed formulation of the Monge problem can be stated as follows:

$$W_p(\mu, \nu) = \inf \left\{ \left( \int_{\Omega \times \Omega} |x - y|^p d\gamma(x, y) \right)^{1/p} : \gamma \in \Pi(\mu, \nu) \right\}.$$

By considering  $\gamma = (\text{id} \times T)_\# \mu$  for any given transport map  $T$  we see that the Kantorovich problem is indeed a relaxation of the Monge problem. It turns out that  $W_p$  is a distance on  $\mathcal{P}(\Omega)$ , which metrizes the weak convergence of measures (recall that we are assuming that  $\Omega$  is bounded). This distance is called the Wasserstein distance of order  $p$ .

Existence and uniqueness results for the optimal mass transportation problem has been proved only very recently. We refer to [8] and [23] for the case  $p = 2$ , to [26] and [19] for the case  $p > 1$  and to [12], [16] and [4] for the case  $p = 1$ . We summarize these results in the following theorem.

**Theorem 2.1** (Existence and uniqueness of an optimal transport map). *Let  $\mu$  and  $\nu$  be probability measures in  $\Omega$  and fix  $p \geq 1$ . We assume that  $\mu$  is absolutely continuous with respect to the Lebesgue measure  $L^d$ . Then  $M_p(\mu, \nu) = W_p(\mu, \nu)$  and there exists an optimal transport map from  $\mu$  to  $\nu$ , which is also unique  $f$ -a.e. if  $p > 1$ .*

**Remark 2.2.** In the particular case when  $\mu \ll L^d$  and  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  any transport map  $T$  is associated to a partition  $(B_i)_{i \in \mathbb{N}}$  of  $\Omega$  in such a way that

$$T(x) = \sum_{i \in \mathbb{N}} y_i \mathbf{1}_{B_i}(x) \quad \text{and} \quad \mu(B_i) = b_i.$$

Conversely, any partition  $(B_i)_{i \in \mathbb{N}}$  of  $\Omega$  satisfying  $\mu(B_i) = b_i$  corresponds to a transport map of the form above.

In the situation of the above remark any transport plan can be written as

$$\gamma = \sum_{i \in \mathbb{N}} (\psi^i \mu) \otimes \delta_{y_i} \quad (2.1)$$

where the functions  $\psi^i : \Omega \rightarrow \mathbb{R}^+$  are such that  $\sum_{i \in \mathbb{N}} \psi^i(x) = 1$  for each  $x \in \Omega$  and  $\int_{\Omega} \psi^i d\mu = b_i$ . With this language  $\gamma$  corresponds to a transport map if and only if the functions  $\psi^i$  have only the values 0 and 1.

Another useful feature of Kantorovich's relaxation is the fact that this new problem admits a dual formulation.

**Theorem 2.3** (Dual formulation). *For  $\mu, \nu \in \mathcal{P}(\Omega)$  the following equality holds:*

$$W_p^p(\mu, \nu) = \sup \left\{ \int_{\Omega} u d\mu + \int_{\Omega} v d\nu : \begin{array}{l} u \in L_{\mu}^1(\Omega), v \in L_{\nu}^1(\Omega) \text{ and} \\ u(x) + v(y) \leq |x - y|^p \text{ for } \mu\text{-a.e. } x \text{ and } \nu\text{-a.e. } y \end{array} \right\}.$$

Moreover, there exists an optimal pair  $(u, v)$  for this dual formulation.

In the particular case when  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  is an atomic measure the dual formulation reads as:

$$W_p^p \left( \mu, \sum_{i \in \mathbb{N}} b_i \delta_{y_i} \right) = \sup \left\{ \int_{\Omega} u d\mu + \sum_{i \in \mathbb{N}} b_i v(y_i) : \begin{array}{l} u \in L_{\mu}^1(\Omega), v \in L_{\nu}^1(\Omega) \text{ and} \\ u(x) + v(y_i) \leq |x - y_i|^p \\ \text{for } \mu\text{-a.e. } x \text{ and every } i \in \mathbb{N} \end{array} \right\}. \quad (2.2)$$

**2.3. A preliminary result on the shape of partitions.** We present in this subsection a result of standard flavour relative to the shape of the sets in the partition determined by the optimal transport of a diffuse measure to an atomic measure.

**Proposition 2.4.** *Let  $f : \Omega \rightarrow \mathbb{R}$  be a nonnegative function such that  $\mu = fL^d \llcorner \Omega$  is a probability measure, and let  $(y_i)_{i \in \mathbb{N}}$  be a sequence of points belonging to  $\Omega$ .*

*(i) Let  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  and  $(B_i)_{i \in \mathbb{N}}$  be a partition of  $\Omega$  such that the map  $T(x) = \sum_{i \in \mathbb{N}} y_i \mathbf{1}_{B_i}(x)$  is an optimal transport map from  $\mu$  to  $\nu$ . Let moreover the pair  $(u, v)$  be any solution of the dual formulation (2.2). Then we have*

$$u(x) = \inf_{i \in \mathbb{N}} \{|x - y_i|^p - v(y_i)\} = \sum_{i \in \mathbb{N}} (|x - y_i|^p - v(y_i)) \mathbf{1}_{B_i}(x) \quad \text{for } f\text{-a.e. } x \in \Omega. \quad (2.3)$$

*(ii) Let  $(B_i)_{i \in \mathbb{N}}$  be a partition of  $\Omega$  and set  $b_i = \int_{B_i} f(x) dx$ ,  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  and  $T(x) = \sum_{i \in \mathbb{N}} y_i \mathbf{1}_{B_i}(x)$ . Let moreover  $u \in L^1_\mu(\Omega)$  and  $v \in L^1_\nu(\Omega)$  be two functions (if any) satisfying condition (2.3). Then we have that  $T$  is optimal for  $M_p(\mu, \nu)$  and that the pair  $(u, v)$  is optimal for the dual formulation (2.2).*

**Remark 2.5.** We easily see that condition (2.3) implies that for every  $i \in \mathbb{N}$  the following equality, intended up to  $f$ -negligible sets, holds:

$$B_i = \{x \in \Omega : |x - y_i|^p - v(y_i) < |x - y_j|^p - v(y_j) \quad \forall j \neq i\}. \quad (2.4)$$

This equality precisely describes the shape of the sets in the partition. For instance, in the case  $p = 1$  the boundaries of the sets  $B_i$  are hyperboloids and in the case  $p = 2$  the cells  $B_i$  are polytopes.

From the above discussion we easily deduce the following corollary, regarding the uniqueness of the optimal transport map from an absolutely continuous measure to an atomic measure, which will be useful in the sequel.

**Corollary 2.6.** *Assume that  $\mu \ll L^d$  and  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  are probability measures in  $\Omega$ . Then the optimal transport map from  $\mu$  to  $\nu$  is unique  $f$ -a.e. even in the case  $p = 1$ .*

*Proof.* The existence of an optimal transport map is ensured by Theorem 2.1. The point is to show that in this case we also have uniqueness. We fix an optimal couple  $(u, v)$  for the dual formulation of the Kantorovich problem (2.2) and we consider an optimal transport map  $T$ . Recalling that, from Remark 2.2, every transport map is associated to a partition of  $\Omega$ , we find a partition  $(B_i)_{i \in \mathbb{N}}$  of  $\Omega$  in such a way that  $T(x) = \sum_{i \in \mathbb{N}} y_i \mathbf{1}_{B_i}(x)$ . But recalling (2.4) we deduce that the sets  $(B_i)_{i \in \mathbb{N}}$  are uniquely determined (up to  $f$ -negligible sets) by  $v$ , thus we conclude that the optimal transport map  $T$  is unique.  $\square$

We close this section proving Proposition 2.4.

*Proof of Proposition 2.4.* (i) Using Theorem 2.3 and the fact that the transport  $T$  and the pair  $(u, v)$  are optimal we deduce that (see Remark 2.2)

$$\int_{\Omega} \sum_{i \in \mathbb{N}} \mathbf{1}_{B_i}(x) |x - y_i|^p f(x) dx = W_p^p(fL^d \llcorner \Omega, \nu) = \int_{\Omega} u(x) f(x) dx + \sum_{i \in \mathbb{N}} b_i v(y_i).$$

Since  $T$  is a transport map from  $\mu$  to  $\nu$  we obtain

$$\int_{\Omega} \sum_{i \in \mathbb{N}} \mathbf{1}_{B_i}(x) |x - y_i|^p f(x) dx = \int_{\Omega} \sum_{i \in \mathbb{N}} \mathbf{1}_{B_i}(x) (u(x) + v(y_i)) f(x) dx.$$

Then, recalling that in the dual formulation (2.2) the pair  $(u, v)$  is subject to the constraint  $u(x) + v(y) \leq |x - y|^p$  for  $\mu$ -a.e.  $x$  and  $\nu$ -a.e.  $y$ , we get

$$\sum_{i \in \mathbb{N}} \mathbf{1}_{B_i}(x) |x - y_i|^p = \sum_{i \in \mathbb{N}} \mathbf{1}_{B_i}(x) (u(x) + v(y_i)) \quad \text{for } f\text{-a.e. } x \in \Omega.$$

This in particular implies

$$u(x) = \sum_{i \in \mathbb{N}} (|x - y_i|^p - v(y_i)) \mathbf{1}_{B_i}(x) \quad \text{for } f\text{-a.e. } x \in \Omega. \quad (2.5)$$

Using again the constraint of the dual formulation (2.2) we immediately deduce (2.3) from (2.5).

(ii) First of all we notice that  $T$  is a transport map from  $f\mathbb{L}^d\mathbf{L}\Omega$  to  $\nu = \sum_{i \in \mathbb{N}} b_i \delta_{y_i}$  and that the pair  $(u, v)$  is admissible for the dual formulation (2.2). Moreover using condition (2.3) we obtain

$$\begin{aligned} \int_{\Omega} |x - T(x)|^p f(x) dx &= \sum_{i \in \mathbb{N}} \int_{B_i} |x - y_i|^p f(x) dx = \sum_{i \in \mathbb{N}} \int_{B_i} [v(y_i) + u(x)] f(x) dx \\ &= \int_{\Omega} v(x) d\nu(y) + \int_{\Omega} u(x) f(x) dx. \end{aligned}$$

Recalling Theorem 2.3 we obtain the validity of (ii).  $\square$

### 3. EXISTENCE, UNIQUENESS AND CHARACTERIZATION OF THE OPTIMUM

**3.1. Formulation of the problem and definition of the optimum.** In this section we consider our problem from a global point of view. The unknown is a partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$ , where each  $A_i$  represents the urban area where customers choosing the service located at  $x_i$  live. The total amount of people living in  $A_i$  is given by  $c_i = \int_{A_i} f(x) dx$ . Every citizen living at  $x$  and using the service located at  $x_i$  is going to spend, in order to be served, a time given by

$$|x - x_i|^p + h_i(c_i).$$

In this expression,  $|x - x_i|^p$  is the transport time and  $h_i(c_i)$  is the time spent waiting in the queue. Consequently, the total time spent by the whole population is given by

$$\sum_{i=1}^k \int_{A_i} [|x - x_i|^p + h_i(c_i)] f(x) dx. \quad (3.1)$$

We want to choose the partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$  in such a way that the quantity in (3.1) is as small as possible; thus we are led to the following minimization problem:

$$\inf_{\substack{(A_i)_{i=1, \dots, k} \\ \text{partition of } \Omega}} \left\{ \sum_{i=1}^k \int_{A_i} \left[ |x - x_i|^p + h_i \left( \int_{A_i} f(x) dx \right) \right] f(x) dx \right\}. \quad (3.2)$$

**Definition 3.1** (Optimum). *We say that a partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$  is an optimum if it is a minimizer for (3.2).*

**3.2. Preliminary considerations.** In this subsection we prove two lemmas which will be used in the following. We denote by  $S$  the unit simplex in  $\mathbb{R}^k$  defined by

$$S = \left\{ c = (c_1, c_2, \dots, c_k) \in \mathbb{R}^k : c_i \geq 0, \sum_{i=1}^k c_i = 1 \right\}.$$

**Lemma 3.2.** *The following equality holds:*

$$\inf (3.2) = \inf_{c \in S} \left\{ W_p^p \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c_i \delta_{x_i} \right) + \sum_{i=1}^k h_i(c_i) c_i \right\}.$$

*Proof.* By applying Theorem 2.1 we have

$$\begin{aligned} & \inf_{c \in S} \left\{ W_p^p \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c_i \delta_{x_i} \right) + \sum_{i=1}^k h_i(c_i) c_i \right\} \\ &= \inf_{c \in S} \left\{ M_p^p \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c_i \delta_{x_i} \right) + \sum_{i=1}^k h_i(c_i) c_i \right\} \quad (3.3) \\ &= \inf_{c \in S} \left\{ \int_{\Omega} |x - T(x)|^p f(x) dx + \sum_{i=1}^k h_i(c_i) c_i : T \text{ transport map from } f \mathbb{L}^d \llcorner \Omega \text{ to } \sum_{i=1}^k c_i \delta_{x_i} \right\}. \end{aligned}$$

But recalling Remark 2.2 we can deduce

$$\inf (3.3) = \inf_{\substack{c \in S \\ (A_i)_{i=1, \dots, k} \\ \text{partition of } \Omega}} \left\{ \sum_{i=1}^k \int_{A_i} |x - x_i|^p f(x) dx + \sum_{i=1}^k h_i(c_i) c_i : \int_{A_i} f(x) dx = c_i \right\} = \inf (3.2),$$

thus we obtain the thesis.  $\square$

**Lemma 3.3.** *The function  $F : S \rightarrow \mathbb{R}$  defined by*

$$F(c_1, \dots, c_k) = W_p^p \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c_i \delta_{x_i} \right)$$

*is continuous and convex.*

*Proof.* The continuity follows directly from the fact that the topology induced by the Wasserstein distance of order  $p$  is the weak convergence of measures, as observed in Section 2.2.

Let us now show the convexity. Let  $(c_1, \dots, c_k)$  and  $(c'_1, \dots, c'_k)$  be two elements of  $S$  and let  $t \in [0, 1]$ . Consider two optimal transport plans

$$\gamma \in \Pi \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c_i \delta_{x_i} \right) \quad \text{and} \quad \gamma' \in \Pi \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k c'_i \delta_{x_i} \right).$$

This implies that

$$t\gamma + (1-t)\gamma' \in \Pi \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k (tc_i + (1-t)c'_i) \delta_{x_i} \right),$$

hence we deduce

$$\begin{aligned} F(t(c_1, \dots, c_k) + (1-t)(c'_1, \dots, c'_k)) &= W_p^p \left( f \mathbb{L}^d \llcorner \Omega, \sum_{i=1}^k (tc_i + (1-t)c'_i) \delta_{x_i} \right) \\ &\leq \int_{\Omega \times \Omega} |x - y|^p d(t\gamma + (1-t)\gamma')(x, y) \\ &= t \left( \int_{\Omega \times \Omega} |x - y|^p d\gamma(x, y) \right) + (1-t) \left( \int_{\Omega \times \Omega} |x - y|^p d\gamma'(x, y) \right) \\ &= tF(c_1, \dots, c_k) + (1-t)F(c'_1, \dots, c'_k). \end{aligned}$$

This precisely means that the map  $F$  is convex.  $\square$

**3.3. Existence and uniqueness of the optimum.** In this subsection we show how to use the reformulation of the minimization problem (3.2) obtained in Lemma 3.2 and the results on optimal mass transportation collected in Section 2.2 to show existence and uniqueness (under suitable assumptions) of an optimum for (3.2). In the following we consider for  $i = 1, \dots, k$  the functions  $\eta_i : [0, 1] \rightarrow [0, +\infty[$  defined by

$$\eta_i(t) = th_i(t). \quad (3.4)$$

**Proposition 3.4.** *Assume that the functions  $h_i$  are lower semi-continuous. Then:*

- (i) *there exists an optimum  $(\hat{A}_i)_{i=1, \dots, k}$  for (3.2);*
- (ii) *if in addition the maps  $\eta_i$  are strictly convex the optimum is unique.*

*Proof.* (i) Recalling the definitions of  $F$  and  $\eta_i$  and using Lemmas 3.2 and 3.3 we have

$$\inf (3.2) = \inf_{c \in S} \left\{ F(c_1, \dots, c_k) + \sum_{i=1}^k \eta_i(c_i) \right\}.$$

The expression to be minimized on the right hand side of the above equality is a lower semi-continuous function defined on the non-empty compact set  $S$ , hence there exists a minimizer  $(\hat{c}_i)_{i=1, \dots, k}$ . By Theorem 2.1 there exists an optimal transport map  $\hat{T}$  from  $f\mathbb{L}^d\mathbb{L}\Omega$  to  $\sum_{i=1}^k \hat{c}_i \delta_{x_i}$ . Thanks to Remark 2.2 the transport map  $\hat{T}$  is associated to a partition  $(\hat{A}_i)_{i=1, \dots, k}$  of  $\Omega$  which satisfies  $\hat{c}_i = \int_{\hat{A}_i} f(x) dx$ . Let us check that this partition is an optimum for (3.2). From the above considerations we deduce

$$\begin{aligned} \inf (3.2) &= \min_{c \in S} \left\{ F(c_1, \dots, c_k) + \sum_{i=1}^k \eta_i(c_i) \right\} = F(\hat{c}_1, \dots, \hat{c}_k) + \sum_{i=1}^k \eta_i(\hat{c}_i) \\ &= W_p^p \left( f\mathbb{L}^d\mathbb{L}\Omega, \sum_{i=1}^k \hat{c}_i \delta_{x_i} \right) + \sum_{i=1}^k \eta_i(\hat{c}_i) \\ &= \sum_{i=1}^k \int_{\hat{A}_i} |x - x_i|^p f(x) dx + \sum_{i=1}^k \eta_i \left( \int_{\hat{A}_i} f(x) dx \right). \end{aligned}$$

This exactly means that  $(\hat{A}_i)_{i=1, \dots, k}$  is an optimum for (3.2).

(ii) Assume by contradiction that we have two different optima  $(\hat{A}_i)_{i=1, \dots, k}$  and  $(\tilde{A}_i)_{i=1, \dots, k}$  for (3.2). We define

$$\hat{c}_i = \int_{\hat{A}_i} f(x) dx \quad \text{and} \quad \tilde{c}_i = \int_{\tilde{A}_i} f(x) dx \quad \text{for each } i = 1, \dots, k.$$

It is immediate to check that  $\hat{T}(x) = \sum_{i=1}^k x_i \mathbf{1}_{\hat{A}_i}(x)$  is an optimal transport map from  $f\mathbb{L}^d\mathbb{L}\Omega$  to  $\sum_{i=1}^k \hat{c}_i \delta_{x_i}$  and that  $\tilde{T}(x) = \sum_{i=1}^k x_i \mathbf{1}_{\tilde{A}_i}(x)$  is an optimal transport map from  $f\mathbb{L}^d\mathbb{L}\Omega$  to  $\sum_{i=1}^k \tilde{c}_i \delta_{x_i}$ . Recalling the fact that the optimal transport map from a diffuse measure to an atomic measure is unique (see Theorem 2.1 for the case  $p > 1$  and Corollary 2.6 for the case  $p = 1$ ) we deduce that  $(\hat{c}_i)_{i=1, \dots, k} \neq (\tilde{c}_i)_{i=1, \dots, k}$ . But using again the result in Lemma 3.2 we deduce that  $(\hat{c}_i)_{i=1, \dots, k}$  and  $(\tilde{c}_i)_{i=1, \dots, k}$  are minimizers of the map  $S \ni (c_1, \dots, c_k) \mapsto F(c_1, \dots, c_k) + \sum_{i=1}^k \eta_i(c_i)$ . However, the strict convexity of the maps  $\eta_i$ , together with the convexity of  $F$  shown in Lemma 3.3, implies the uniqueness of the minimizer of this map, and from this contradiction we deduce the thesis.  $\square$

**3.4. Characterization of the optimum.** We say that a function  $\phi : ]0, 1] \rightarrow [0, +\infty[$  is differentiable if it is differentiable in the usual sense in the open interval  $]0, 1[$  and has a left derivative in the point 1. We also use the convention  $0h'_i(0) = 0$ .

**Proposition 3.5** (Necessary optimality condition). *Let  $(A_i)_{i=1,\dots,k}$  be an optimum for (3.2). Assume that  $h_i$  are differentiable in  $]0, 1]$  and continuous in 0. Then for all  $i = 1, \dots, k$  the following holds:*

$$\begin{cases} A_i = \left\{ x \in \Omega : |x - x_i|^p + h_i(c_i) + c_i h'_i(c_i) < |x - x_j|^p + h_j(c_j) + c_j h'_j(c_j) \quad \forall j \neq i \right\} \\ c_i = \int_{A_i} f(x) dx, \end{cases} \quad (3.5)$$

where as usual equalities between sets are intended up to  $f$ -negligible sets.

*Proof.* Consider a partition  $(A_r)_{r=1,\dots,k}$  which is an optimum for (3.2) and take an  $\mathbb{L}^d$ -negligible set  $N \subset \Omega$  such that for every  $x \notin N$

- (i)  $x$  is a Lebesgue point of  $f$ ;
- (ii) if  $x \in A_r$ , then  $x$  is a point of density 1 in  $A_r$ .

Fix now two indices  $i$  and  $j$  in  $\{1, \dots, k\}$  and consider a point  $x_0 \in A_i \setminus N$ . For every  $\epsilon > 0$  let the partition  $(\tilde{A}_r)_{r=1,\dots,k}$  be defined by

$$\begin{cases} \tilde{A}_i = A_i \setminus B_\epsilon(x_0) \\ \tilde{A}_j = A_j \cup (A_i \cap B_\epsilon(x_0)) \\ \tilde{A}_r = A_r \text{ for all } r \in \{1, \dots, k\} \setminus \{i, j\}. \end{cases}$$

Let  $c_\epsilon = \int_{A_i \cap B_\epsilon(x_0)} f(x) dx$ . By the optimality of the partition  $(A_r)_{r=1,\dots,k}$ , comparing its total cost with the total cost of the partition  $(\tilde{A}_r)_{r=1,\dots,k}$  we deduce

$$\begin{aligned} \sum_{r=1}^k \int_{A_r} |x - x_r|^p f(x) dx + \sum_{r=1}^k c_r h_r(c_r) &\leq \sum_{\substack{r=1,\dots,k \\ r \neq i,j}} \int_{A_r} |x - x_r|^p f(x) dx \\ &+ \sum_{\substack{r=1,\dots,k \\ r \neq i,j}} c_r h_r(c_r) + \int_{A_i \setminus B_\epsilon(x_0)} |x - x_i|^p f(x) dx + (c_i - c_\epsilon) h_i(c_i - c_\epsilon) \\ &+ \int_{A_j \cup (A_i \cap B_\epsilon(x_0))} |x - x_j|^p f(x) dx + (c_j + c_\epsilon) h_j(c_j + c_\epsilon). \end{aligned}$$

This leads to

$$\begin{aligned} \int_{A_i \cap B_\epsilon(x_0)} |x - x_i|^p f(x) dx + c_i h_i(c_i) + c_j h_j(c_j) \\ \leq (c_i - c_\epsilon) h_i(c_i - c_\epsilon) + \int_{A_i \cap B_\epsilon(x_0)} |x - x_j|^p f(x) dx + (c_j + c_\epsilon) h_j(c_j + c_\epsilon). \end{aligned}$$

We now divide this expression by  $\omega_d \epsilon^d$ , where  $\omega_d$  is the volume of the unit ball in  $\mathbb{R}^d$ . We obtain

$$\begin{aligned} \frac{1}{\omega_d \epsilon^d} \int_{A_i \cap B_\epsilon(x_0)} |x - x_i|^p f(x) dx + \frac{c_i}{\omega_d \epsilon^d} [h_i(c_i) - h_i(c_i - c_\epsilon)] + \frac{c_\epsilon}{\omega_d \epsilon^d} h_i(c_i - c_\epsilon) \\ \leq \frac{1}{\omega_d \epsilon^d} \int_{A_i \cap B_\epsilon(x_0)} |x - x_j|^p f(x) dx + \frac{c_j}{\omega_d \epsilon^d} [h_j(c_j + c_\epsilon) - h_j(c_j)] + \frac{c_\epsilon}{\omega_d \epsilon^d} h_j(c_j + c_\epsilon). \end{aligned}$$

Letting  $\varepsilon \rightarrow 0$  and recalling that  $x_0$  satisfies assumptions (i) and (ii) we easily obtain

$$|x_0 - x_i|^p f(x_0) + f(x_0) c_i h'_i(c_i) + f(x_0) h_i(c_i) \leq |x_0 - x_j|^p f(x_0) + f(x_0) c_j h'_j(c_j) + f(x_0) h_j(c_j),$$

that is the desired thesis.  $\square$

The following lemma will imply that condition (3.5) is in fact sufficient, under reasonable assumptions.

**Lemma 3.6.** *Assume that the functions  $h_i$  are differentiable in  $]0, 1]$  and continuous in 0 and that  $\eta_i$  are convex. Then there exists at most one partition  $(A_i)_{i=1, \dots, k}$  for which (3.5) holds.*

We notice that the convexity assumption on  $\eta_i$  is satisfied for instance if  $h_i(t) = t^q$  with  $0 < q < 1$ , which are natural concave cost functions.

*Proof of Lemma 3.6.* Assume by contradiction that there exist two partitions  $(A_i)_{i=1, \dots, k}$  and  $(\hat{A}_i)_{i=1, \dots, k}$  satisfying (3.5). We set  $c_i = \int_{A_i} f(x) dx$  and  $\hat{c}_i = \int_{\hat{A}_i} f(x) dx$  and we consider the set  $I = \{i \in \{1, \dots, k\} : c_i < \hat{c}_i\}$ . Assume that  $I$  is nonempty. Recalling the definition of the functions  $\eta_i$  and comparing with (3.5) we obtain

$$\bigcup_{i \in I} A_i = \left\{ x \in \Omega : \min_{i \in I} |x - x_i|^p + \eta'_i(c_i) < \min_{j \notin I} |x - x_j|^p + \eta'_j(c_j) \right\}$$

and

$$\bigcup_{i \in I} \hat{A}_i = \left\{ x \in \Omega : \min_{i \in I} |x - x_i|^p + \eta'_i(\hat{c}_i) < \min_{j \notin I} |x - x_j|^p + \eta'_j(\hat{c}_j) \right\}.$$

Since we are assuming that the functions  $\eta_i$  are convex, we obtain that  $\eta'_i(c_i) \leq \eta'_i(\hat{c}_i)$  for every  $i \in I$  and that  $\eta'_j(c_j) \geq \eta'_j(\hat{c}_j)$  for every  $j \notin I$ . This immediately implies that  $\bigcup_{i \in I} \hat{A}_i \subseteq \bigcup_{i \in I} A_i$ , thus

$$\int_{\bigcup_{i \in I} \hat{A}_i} f(x) dx \leq \int_{\bigcup_{i \in I} A_i} f(x) dx. \quad (3.6)$$

But recalling the definition of  $c_i$  and  $\hat{c}_i$  and of the set  $I$  we have

$$\int_{\bigcup_{i \in I} A_i} f(x) dx = \sum_{i \in I} c_i < \sum_{i \in I} \hat{c}_i = \int_{\bigcup_{i \in I} \hat{A}_i} f(x) dx,$$

and this is in contradiction with (3.6). We deduce that the set  $I$  is empty, which means that  $c_i \geq \hat{c}_i$  for every  $i = 1, \dots, k$ . By symmetry the opposite inequality is also true, thus we obtain that  $c_i = \hat{c}_i$  for every  $i = 1, \dots, k$ . But going back to (3.5) we immediately deduce that  $A_i = \hat{A}_i$  for every  $i = 1, \dots, k$ .  $\square$

**Proposition 3.7** (Sufficient optimal condition). *Let  $(A_i)_{i=1, \dots, k}$  be a partition of  $\Omega$  which satisfies (3.5). Assume that the functions  $h_i$  are differentiable in  $]0, 1]$  and continuous in 0 and that  $\eta_i$  are convex. Then  $(A_i)_{i=1, \dots, k}$  is the unique optimum for (3.2).*

*Proof.* It is a direct consequence of Propositions 3.4 and 3.5 and of Lemma 3.6.  $\square$

**3.5. Summary of the results on the optimum.** We close this section presenting a summary of the results relative to the optimum for (3.2). Remember that the functions  $\eta_i$  have been defined in (3.4).

- If  $h_i$  are lower semi-continuous, then there exists an optimum for (3.2).
- If  $h_i$  are lower semi-continuous and  $\eta_i$  are strictly convex, then there exists a unique optimum for (3.2).

- If  $h_i$  are differentiable in  $]0, 1]$  and continuous in 0 and  $\eta_i$  are convex, then there exists a unique optimum for (3.2).
- If  $h_i$  are differentiable and continuous in 0, then (3.5) is a necessary optimality condition.
- If  $h_i$  are differentiable in  $]0, 1]$  and continuous in 0 and  $\eta_i$  are convex, then (3.5) is a necessary and sufficient optimality condition.

We remark in passing what follows: in the general setting of the Monge problem, the “distance”  $|y - x|^p$  is replaced by any lower semi-continuous function  $c(x, y)$ , usually called cost function. But in fact we have used only two features of the particular choice  $c(x, y) = |y - x|^p$ , namely, the existence and uniqueness of an optimal transport map and the fact that for any  $y \in \Omega$  the level sets of  $c(\cdot, y)$  are negligible. Hence our results can be extended for general costs  $c(x, y)$  for which both properties hold true, such as  $c(x, y) = |y - x|^p$  with  $0 < p < 1$  (see [19]) or  $c(x, y) = \|y - x\|$  for different kinds of norm (see [12] and [3]).

#### 4. EXISTENCE, UNIQUENESS AND CHARACTERIZATION OF THE EQUILIBRIUM

**4.1. Definition of the equilibrium.** In this section we start to consider the situation from the point of view of the single citizen. Assume for the moment that the sets  $A_i$ , and so the quantities  $c_i$ , are given: the single citizen is probably not interested whether the partition  $(A_i)_{i=1, \dots, k}$  is optimal in the global sense that we discussed up to now. Indeed, it is quite convincing that he does not even know the sets  $A_i$ , nor the quantities  $c_i$ , nor the functions  $h_i$ : what is meaningful, is that he only knows the quantities  $h_i(c_i)$ , i.e. the queue times in the various services. Then, of course, a citizen living at  $x \in \Omega$  and going to  $x_i$  shall be “satisfied” if and only if

$$|x - x_i|^p + h_i(c_i) = \min_{j=1, \dots, k} |x - x_j|^p + h_j(c_j).$$

In the context of game theory this is precisely a Nash equilibrium (see Section 12.3 of [5]): each player is satisfied in the sense that his strategy (i.e. the choice of the service) is the best possible, once the behaviour of the other players (in this case, the sets  $A_i$ ) is fixed. We give the following definition.

**Definition 4.1.** *We say that a partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$  is an equilibrium if for every  $i = 1, \dots, k$  the following condition holds:*

$$\begin{cases} A_i = \left\{ x \in \Omega : |x - x_i|^p + h_i(c_i) < |x - x_j|^p + h_j(c_j) \text{ for every } j \neq i \right\} \\ c_i = \int_{A_i} f(x) dx. \end{cases} \quad (4.1)$$

We remark that this type of equilibrium is non-cooperative: each citizen chooses the service just by himself, without collaborating with the other citizens in order to choose a “better” global strategy. In Theorem 4.4 and Proposition 4.5 we show that there exists an equilibrium, under fairly general assumptions: we remark that this is not a trivial result, since in general game theory it is typically a difficult task to show existence of a Nash equilibrium.

**4.2. Existence and uniqueness of the equilibrium.** In this subsection we show how it is possible to formulate an auxiliary problem in such a way that the optimum for this new problem corresponds to the equilibrium for the original problem. Thus we easily deduce some existence and uniqueness results for the equilibrium, relying on the results presented in the previous section.

Assume that the functions  $h_i$  are continuous. We introduce for  $i = 1, \dots, k$  the functions  $g_i : [0, 1] \rightarrow [0, +\infty[$  defined by

$$g_i(t) = \begin{cases} \frac{1}{t} \int_0^t h_i(s) ds & \text{if } 0 < t \leq 1 \\ h_i(0) & \text{if } t = 0. \end{cases}$$

It is immediate that the functions  $g_i$  are continuous and that  $t \mapsto tg_i(t)$  are differentiable everywhere in  $[0, 1]$  with derivative  $h_i(t)$ .

**Proposition 4.2.** *Assume that the functions  $h_i$  are continuous. Every partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$  which is a minimizer of the problem*

$$\inf_{\substack{(A_i)_{i=1, \dots, k} \\ \text{partition of } \Omega}} \left\{ \sum_{i=1}^k \int_{A_i} \left[ |x - x_i|^p + g_i \left( \int_{A_i} f(x) dx \right) \right] f(x) dx \right\} \quad (4.2)$$

*is an equilibrium for the original problem. If in addition the functions  $h_i$  are non-decreasing then every equilibrium for the original problem is a minimizer of (4.2).*

*Proof.* Consider a minimizer  $(A_i)_{i=1, \dots, k}$  of (4.2). Since the functions  $g_i$  are continuous in 0 and the functions  $tg_i(t)$  are differentiable in  $[0, 1]$  we can apply Proposition 3.5 to deduce that  $(A_i)_{i=1, \dots, k}$  satisfies the optimality condition (3.5), but in view of the definition of  $g_i$  this condition is precisely (4.1): recall that  $(tg_i(t))' = h_i(t)$ . Thus  $(A_i)_{i=1, \dots, k}$  is an equilibrium for the original problem. If  $h_i$  is non-decreasing for all  $i = 1, \dots, k$ , we immediately obtain the convexity of the functions  $t \mapsto tg_i(t)$ , thus we can apply Proposition 3.7, obtaining precisely that every equilibrium for the original problem is a minimizer of (4.2).  $\square$

**Remark 4.3.** Arguing similarly to the previous proof, we can also show that, under the assumption that the maps  $\eta_i$  defined in (3.4) are differentiable in  $[0, 1]$ , any optimum for (3.2) is an equilibrium for the problem with queue functions

$$\tilde{h}_i(t) = h_i(t) + th_i'(t) \quad i = 1, \dots, k.$$

Using the correspondence given by the previous proposition, it is now easy to show the following theorem.

**Theorem 4.4.** *Assume that the functions  $h_i$  are continuous. Then there exists an equilibrium. If in addition the functions  $h_i$  are non-decreasing then the equilibrium is unique.*

*Proof.* Under the assumptions of the theorem we have that the functions  $t \mapsto tg_i(t)$  are lower semi-continuous, thus applying Proposition 3.4 we obtain the existence of a minimizer of (4.2), and using the result of Proposition 4.2 we deduce that this minimizer is an equilibrium for the original problem. If the functions  $h_i$  are non-decreasing we apply again Proposition 4.2, obtaining that every equilibrium for the original problem is a minimizer of (4.2). But since  $h_i$  are non-decreasing we also deduce that the maps  $t \mapsto tg_i(t)$  are convex, thus using Proposition 3.7 we deduce that (4.2) has a unique solution, and this concludes the proof.  $\square$

**4.3. A direct proof in the case  $k = 2$ .** We now give an alternative and more direct proof of the above result in the particular case  $k = 2$ . This has also the advantage of introducing

some ideas and notation that will be important in the dynamical analysis of Section 5. In what follows, we will make extensively use of the following definitions:

$$\tau(x) = |x - x_1|^p - |x - x_2|^p, \quad m(t) = \int_{\{x : \tau(x) < t\}} f(x) dx. \quad (4.3)$$

The condition in (4.1) defining the equilibrium can be read in this context as follows: *a partition  $(A_1, A_2)$  of  $\Omega$  is an equilibrium when there exists  $\bar{t} \in \mathbb{R}$  such that:*

$$A_1 = \{x \in \Omega : \tau(x) < \bar{t}\}, \quad A_2 = \{x \in \Omega : \tau(x) > \bar{t}\} \quad (4.4)$$

$$h_2(1 - m(\bar{t})) - h_1(m(\bar{t})) = \bar{t}. \quad (4.5)$$

**Proposition 4.5.** *Assume that the functions  $h_1$  and  $h_2$  are continuous and non-decreasing. Then there exists a unique equilibrium  $(A_1, A_2)$ .*

*Proof.* Let us consider the map

$$U(t) = t - h_2(1 - m(t)) + h_1(m(t)).$$

It is a continuous and strictly increasing map. Notice that for  $t \geq \sup_{\Omega} \tau$  we have  $m(t) = 1$  and  $U(t) = t - h_2(0) + h_1(1)$  so that  $\lim_{t \rightarrow +\infty} U(t) = +\infty$ . In the same way for  $t \leq \inf_{\Omega} \tau$  we have  $m(t) = 0$  and  $U(t) = t - h_2(1) + h_1(0)$ , thus  $\lim_{t \rightarrow -\infty} U(t) = -\infty$ . By consequence, there exists a unique  $\bar{t}$  such that  $U(\bar{t}) = 0$ , that is, (4.5) holds. Then the partition  $(A_1, A_2)$  associated to  $\bar{t}$  as in (4.4) is an equilibrium and this is unique.  $\square$

**4.4. A comparison between optimum and equilibrium.** Recall that the optimality condition for the optimum obtained in Proposition 3.5, read with the notation of the previous subsection, was

$$h_2(1 - m(\hat{t})) + (1 - m(\hat{t})) h'_2(1 - m(\hat{t})) - h_1(m(\hat{t})) - m(\hat{t}) h'_1(m(\hat{t})) = \hat{t}.$$

Thus a comparison with (4.5) shows how deep is the difference between the two conditions of optimality. In this subsection we give two explicit examples in which we see in a qualitative way some differences between the optimum and the equilibrium.

**Example 4.6.** On a beach represented by  $\Omega = [0, 1]$  there are two ice-cream shops at coordinates  $x_1 = 1/4$  and  $x_2 = 3/4$ . Suppose there are less employees in the second shop, so that  $h_1(t) = t$  and  $h_2(t) = (1 + \varepsilon)t$ . Assume that the costumers are uniformly distributed on the beach, that is  $f \equiv 1$ , and take  $p = 2$ .

In this case the optimum  $(A_1, A_2)$  is given by

$$A_1 = [0, \lambda_{\varepsilon}^{\text{opt}}[, \quad A_2 = ]\lambda_{\varepsilon}^{\text{opt}}, 1] \quad \text{with } \lambda_{\varepsilon}^{\text{opt}} = \frac{1}{2} + \frac{\varepsilon}{5 + 2\varepsilon},$$

whereas the equilibrium  $(B_1, B_2)$  is

$$B_1 = [0, \lambda_{\varepsilon}^{\text{eq}}[, \quad B_2 = ]\lambda_{\varepsilon}^{\text{eq}}, 1] \quad \text{with } \lambda_{\varepsilon}^{\text{eq}} = \frac{1}{2} + \frac{\varepsilon}{6 + 2\varepsilon} \leq \lambda_{\varepsilon}^{\text{opt}}.$$

The costs for the customers in the two cases are given by:

cost on the optimum	cost on the equilibrium	position
$ x - 1/4 ^2 + \lambda_\varepsilon^{\text{opt}}$	$ x - 1/4 ^2 + \lambda_\varepsilon^{\text{eq}}$	$0 \leq x \leq \lambda_\varepsilon^{\text{eq}}$
$ x - 1/4 ^2 + \lambda_\varepsilon^{\text{opt}}$	$ x - 3/4 ^2 + (1 + \varepsilon)(1 - \lambda_\varepsilon^{\text{eq}})$	$\lambda_\varepsilon^{\text{eq}} \leq x \leq \lambda_\varepsilon^{\text{opt}}$
$ x - 3/4 ^2 + (1 + \varepsilon)(1 - \lambda_\varepsilon^{\text{opt}})$	$ x - 3/4 ^2 + (1 + \varepsilon)(1 - \lambda_\varepsilon^{\text{eq}})$	$\lambda_\varepsilon^{\text{opt}} \leq x \leq 1$

Note that, in the case  $\varepsilon = 0$ , the situations of optimum and equilibrium coincides and we have  $A_1 = B_1 = [0, 1/2[$  and  $A_2 = B_2 = ]1/2, 1]$ . In the case  $\varepsilon > 0$ , despite the global cost is minimized by the partition  $(A_1, A_2)$ , in the equilibrium situation more than one half of the costumers pay less than in the optimum situation.

**Example 4.7.** Let us now take  $\Omega = [0, 1]$ ,  $f \equiv 1$ ,  $x_1 = 0$ ,  $x_2 = 1$ ,  $p = 1$  and

$$h_1(s) \equiv 100 \quad \text{and} \quad h_2(s) = \begin{cases} 0 & \text{for } 0 \leq s \leq 0.999 \\ 1 & \text{for } 0.999 < s \leq 1. \end{cases}$$

It is clear that the unique equilibrium is  $A_1 = \emptyset$  and  $A_2 = [0, 1]$ , whereas the optimum is obtained with  $A_1 = [0, 0.001[$  and  $A_2 = ]0.001, 1]$ . Notice that the optimum is very unfair for costumers living in  $A_1$ , who pay  $100 + x$ , whereas the other costumers just pay the distance from 1. Note that the result would have the same features with  $h_2$  equal to 0 on  $[0, 0.998]$ , equal to 1 on  $[0.999, 1]$  and increasing smoothly from 0 to 1 on  $]0.998, 0.999[$ .

**4.5. A comparison with Pareto optimum.** The examples given in the previous subsection show that in some cases the global optimum can really be not convenient for some citizens; in particular, in Example 4.7 the optimum is better than the equilibrium for 99.9% of the citizens, even though it is much worse for the remaining 0.1%. In some classical games (like the well-known prisoner's dilemma, see for instance Section 7.8 of [5]) the equilibrium is in fact a bad strategy for *all* players, in the sense that there is a situation which is better for everybody.

On the contrary, we are going to see that in our model this can not happen. Even more, given an equilibrium, we show that it is not possible to find a situation in which every citizen spend less time. More precisely, and introducing a further notion from game theory, we are going to see that every equilibrium is a Pareto optimum for the problem. This means that, in our situation, starting from the non-cooperative Nash equilibrium it is not possible to lower the costs of all the citizens by a cooperation in the choice of the services.

Let us introduce the individual cost function

$$C(x, (B_i)_i) = \sum_{i=1}^k \left[ |x - x_i|^p + h_i \left( \int_{B_i} f(x) dx \right) \right] \mathbf{1}_{B_i}(x), \quad (4.6)$$

defined for  $x \in \Omega$  and  $(B_i)_{i=1, \dots, k}$  partition of  $\Omega$ . We say that a partition  $(A_i)_{i=1, \dots, k}$  of  $\Omega$  is a *Pareto optimum* (see also Section 12.5 of [5]) if there exists no partition  $(B_i)_{i=1, \dots, k}$  of  $\Omega$  satisfying

$$C(x, (B_i)_i) \leq C(x, (A_i)_i) \quad \text{for } f\text{-a.e. } x \in \Omega$$

with strict inequality in a set of strictly positive measure. With this definition of  $C$  the following equivalence is immediate:

$$(A_i)_{i=1,\dots,k} \text{ is an equilibrium } \iff C(x, (A_i)_i) = \min_{j=1,\dots,k} \left\{ |x - x_j|^p + h_j \left( \int_{A_j} f(x) dx \right) \right\}. \quad (4.7)$$

We establish now the following result of independent interest, which makes a connection between the equilibrium and the optimal transport map from  $\mu = fL^d \llcorner \Omega$  to the atomic measure concentrated on the points  $x_1, \dots, x_k$ .

**Proposition 4.8** (Link between equilibrium and optimal transportation). *Let  $(A_i)_{i=1,\dots,k}$  be a partition of  $\Omega$ ,  $\mu = fL^d \llcorner \Omega$ ,  $c_i = \int_{A_i} f$ ,  $\nu = \sum_{i=1}^k c_i \delta_{x_i}$  and  $T = \sum_{i=1}^k x_i \mathbf{1}_{A_i}$ ; let moreover  $u(x) = C(x, (A_i)_i)$  and  $v(x_i) = -h_i(c_i)$ . Then the couple  $(u, v)$  is optimal for the dual problem (2.2) if and only if  $(A_i)_{i=1,\dots,k}$  is an equilibrium. Moreover, in this case  $T$  is an optimal transport map between  $\mu$  and  $\nu$ .*

*Proof.* First of all, one can notice that the cost of the transport  $T$  equals  $\int_{\Omega} u d\mu + \sum_i c_i v(x_i)$ : indeed, one has

$$\begin{aligned} \int_{\Omega} |x - T(x)|^p d\mu &= \sum_{i=1}^k \int_{A_i} |x - x_i|^p f(x) dx = \sum_{i=1}^k \int_{A_i} \left[ |x - x_i|^p + h_i(c_i) \right] f(x) dx + \sum_{i=1}^k c_i v(x_i) \\ &= \int_{\Omega} u(x) d\mu(x) + \int_{\Omega} v(x) d\nu(x). \end{aligned}$$

Thanks to Theorem 2.3 we deduce that, if  $(u, v)$  is admissible for the dual problem, then  $T$  is an optimal transport map and  $(u, v)$  is optimal for the dual problem.

On the other hand, the property (4.7) exactly means that  $(u, v)$  is admissible for the dual problem if and only if  $(A_i)_{i=1,\dots,k}$  is an equilibrium. These two considerations give the thesis.  $\square$

Notice that, in the situation of the proposition above, it may happen that  $T$  is an optimal map but  $(u, v)$  is *not* optimal for the dual problem and  $(A_i)_{i=1,\dots,k}$  is *not* an equilibrium.

We now use Proposition 4.8 to prove that an equilibrium is a Pareto optimum.

**Proposition 4.9.** *Assume that the maps  $h_i$  are strictly increasing and let  $(A_i)_{i=1,\dots,k}$  be an equilibrium. Let  $(B_i)_{i=1,\dots,k}$  be a partition of  $\Omega$  such that*

$$C(x, (B_i)_i) \leq C(x, (A_i)_i) \quad \text{for } f\text{-a.e. } x \in \Omega. \quad (4.8)$$

*Then the partitions  $(A_i)_{i=1,\dots,k}$  and  $(B_i)_{i=1,\dots,k}$  coincide.*

*Proof.* We first show that for every  $i = 1, \dots, k$  the equality  $\int_{A_i} f(x) dx = \int_{B_i} f(x) dx$  holds. Indeed, if this were not true, we could find  $j \in \{1, \dots, k\}$  such that  $\int_{A_j} f(x) dx < \int_{B_j} f(x) dx$ . Considering a point  $x \in B_j$  such that (4.8) holds we have

$$\begin{aligned} |x - x_j|^p + h_j \left( \int_{B_j} f(x) dx \right) &= C(x, (B_i)_i) \leq C(x, (A_i)_i) \\ &= \min_{i=1,\dots,k} \left\{ |x - x_i|^p + h_i \left( \int_{A_i} f(x) dx \right) \right\} \\ &\leq |x - x_j|^p + h_j \left( \int_{A_j} f(x) dx \right). \end{aligned}$$

But this leads to a contradiction: indeed,  $h_j$  being strictly increasing, we should have  $h_j \left( \int_{A_j} f(x) dx \right) < h_j \left( \int_{B_j} f(x) dx \right)$ . This shows as claimed that  $\int_{A_i} f(x) dx = \int_{B_i} f(x) dx$  for every  $i = 1, \dots, k$ .

Since  $(A_i)_{i=1, \dots, k}$  is an equilibrium, using Proposition 4.8 we obtain that the map  $T = \sum_{i=1}^k x_i \mathbf{1}_{A_i}$  is an optimal transport map from  $\mu$  to  $\sum_{i=1}^k \left( \int_{A_i} f(x) dx \right) \delta_{x_i}$ . Using this fact and applying the assumption (4.8) we deduce

$$\begin{aligned} \sum_{i=1}^k \int_{B_i} |x - x_i|^p f(x) dx &\leq \sum_{i=1}^k \int_{A_i} |x - x_i|^p f(x) dx \\ &= W_p \left( \mu, \sum_{i=1}^k \left( \int_{A_i} f(x) dx \right) \delta_{x_i} \right) = W_p \left( \mu, \sum_{i=1}^k \left( \int_{B_i} f(x) dx \right) \delta_{x_i} \right). \end{aligned}$$

Hence the map  $S = \sum_{i=1}^k x_i \mathbf{1}_{B_i}$  is an optimal transport map from  $\mu$  to  $\sum_{i=1}^k \left( \int_{B_i} f(x) dx \right) \delta_{x_i} = \sum_{i=1}^k \left( \int_{A_i} f(x) dx \right) \delta_{x_i}$ . By the uniqueness results of Theorem 2.1 and Corollary 2.6 we conclude.  $\square$

**Corollary 4.10.** *If the maps  $h_i$  are strictly increasing, then every equilibrium is a Pareto optimum.*

## 5. EVOLUTION DYNAMIC AND CONVERGENCE TO THE EQUILIBRIUM

In this section we discuss a dynamical formulation of the problem, in which every single day each citizen decides where to go, using the knowledge of what happened in the previous days and trying to make a “smart” choice. We set the problem and discuss under which assumptions the situation converges to the (unique) individual equilibrium, giving also counterexamples to show that this is not always the case. In the first subsection we consider the standard evolution, where every day one thinks only to the previous day. In the second subsection we study the evolution with prudence, where the citizens again remember only the previous day, but are more careful and try to avoid changing idea too often. In the last subsection we describe the case of the evolution with memory, where each day the citizens remember more previous days. In all this section, to keep the discussion as simple as possible, we consider the case  $k = 2$  when there are only two services located at  $x_1$  and  $x_2$ ; however, the general case requires more care with multiple indices but no really new ideas.

Recall that, thanks to Proposition 4.5, we know the existence of a unique individual equilibrium under the assumption that the functions  $h_1$  and  $h_2$  are continuous and non-decreasing.

**5.1. Standard evolution.** We face now the question whether or not the situation can naturally evolve toward this equilibrium. By “naturally” we mean that we try to model the evolution in time, in which the citizens do not know the situation in its complete complexity, but they just see how long is the queue and decide day by day how to behave. In other words, we introduce a simple scheme which models the everyday choice of people, and we study the possible convergence under this model. The idea is very easy: at each day, each citizen decides freely whether to go to  $x_1$  or to  $x_2$ . Moreover, at the end of every day he discovers what has happened in that day, that is, he discovers  $h_1(c_1)$  and  $h_2(c_2)$ : this will clearly affect the choice of the following day. More precisely, for each  $j \in \mathbb{N}$  we have the function  $\psi_j : \Omega \rightarrow \{0, 1\}$ , where  $\psi_j(x) = 0$  (resp.  $\psi_j(x) = 1$ ) means that at the day  $j$  the citizen living at  $x$  goes to  $x_1$  (resp. to  $x_2$ ); note that,

in the language of (2.1), the function  $\psi_j$  corresponds to  $\psi^2$  at time-step  $j$ , while  $\psi^1$  is simply  $1 - \psi^2$ . We also set  $m_j \in [0, 1]$  the mass of the people that, at the day  $j$ , decide to go to  $x_1$ , i.e.

$$m_j := \int_{\Omega} (1 - \psi_j(x)) f(x) dx. \quad (5.1)$$

The evolution of the problem is the following: we fix any function  $\psi_0 : \Omega \rightarrow \{0, 1\}$ , and consider it as a data of the problem; that is, we leave the citizens free to decide by chance what to do at the day 0. Of course, the most meaningful choice would be

$$\psi_0(x) = \begin{cases} 0 & \text{if } \tau(x) < 0 \\ 1 & \text{otherwise,} \end{cases} \quad (5.2)$$

recalling the definitions of  $\tau$  and  $m$  given previously

$$\tau(x) = |x - x_1|^p - |x - x_2|^p, \quad m(t) = \int_{\{x : \tau(x) < t\}} f(x) dx. \quad (4.3)$$

Indeed, this choice of  $\psi_0$  means that in the first day each one simply goes to the closest of the two service places, which makes sense since he does not have any information about queues. However, our results also hold with any other function  $\psi_0$ . Given now any  $j \in \mathbb{N}$ , and given the function  $\psi_j$ , the function  $\psi_{j+1}$  is determined as follows:

$$\psi_{j+1}(x) = \begin{cases} 0 & \text{if } \tau(x) < h_2(1 - m_j) - h_1(m_j) \\ 1 & \text{otherwise.} \end{cases} \quad (5.3)$$

The meaning of this model is extremely clear: at the day  $j + 1$ , each citizen decides where to go assuming that the quantity of people going to  $x_1$  and  $x_2$  will be the same as in the day  $j$ . The first property that we can easily notice is the following: there exists a sequence  $(t_j)_{j \geq 1}$  such that for any  $j$  one has

$$\psi_j(x) = \begin{cases} 0 & \text{if } \tau(x) < t_j \\ 1 & \text{otherwise;} \end{cases}$$

this is immediate by the definition for any  $j \geq 1$ . Being  $\psi_0$  free, we can not say the same for  $j = 0$ , of course: however, in the case of the “meaningful” choice (5.2), the above formula holds also for  $j = 0$  with  $t_0 = 0$ . Notice that

$$t_{j+1} = h_2(1 - m_j) - h_1(m_j); \quad (5.4)$$

as a consequence, we understand that everything in the choices of the citizen living at  $x$  depends only on the value of  $\tau(x)$ . This also implies, for  $j \geq 1$ , that

$$m_j = m(t_j).$$

What we want to know is if the situation converges to the individual optimum; that is, if  $\psi_j$  converges to the function  $\bar{\psi}$  defined as

$$\bar{\psi}(x) = \begin{cases} 0 & \text{if } x \in A_1 \\ 1 & \text{if } x \in A_2, \end{cases}$$

where  $(A_1, A_2)$  is the equilibrium partition given by Proposition 4.5. Equivalently, we want to understand if  $t_j$  converges to  $\bar{t}$  where, according to Proposition 4.5,  $\bar{t}$  is the unique real number such that

$$\bar{t} = h_2(1 - m(\bar{t})) - h_1(m(\bar{t})).$$

A first answer to this question is given below. We shall consider the function  $G : \mathbb{R} \rightarrow \mathbb{R}$  defined by

$$G(t) := h_2(1 - m(t)) - h_1(m(t)). \quad (5.5)$$

Notice that the function  $G$  has a very simple meaning, namely, it is  $t_{j+1} = G(t_j)$ ; moreover  $\bar{t}$  is uniquely determined by the fact that  $G(\bar{t}) = \bar{t}$ . We then prove the following result which clarifies what happens for the standard evolution.

**Proposition 5.1.** *Assume that  $h_1$  and  $h_2$  are continuous and non-decreasing, as well as that*

$$|G(s) - G(t)| < |s - t| \quad (5.6)$$

*for all  $s \neq t \in \mathbb{R}$ . Then  $t_j \rightarrow \bar{t}$  for  $j \rightarrow +\infty$ , hence also  $\psi_j \rightarrow \bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

*Proof.* The evolution of the sequence  $j \mapsto t_j$  is shown in the Figure 1: notice that since  $h_1$  and  $h_2$  are non-decreasing then the function  $G$  is non-increasing, hence the equilibrium point  $\bar{t} = G(\bar{t})$  is unique. Let us assume by symmetry that, as in the figure,  $t_1 \leq \bar{t}$ : then, since  $G$  is a

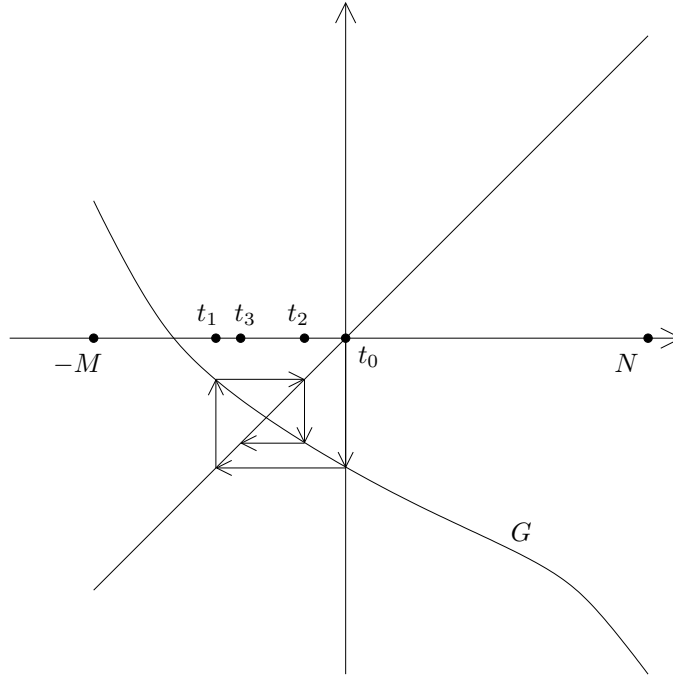


FIGURE 1. The situation in Proposition 5.1

non-increasing function and  $G(\bar{t}) = \bar{t}$ , we have immediately that  $t_n \geq \bar{t}$  for all even  $n$  and  $t_n \leq \bar{t}$  for all odd  $n$ . Therefore,  $t_{2n} \searrow t^+$  and  $t_{2n+1} \nearrow t^-$  with  $t^- \leq \bar{t} \leq t^+$ . Moreover, by continuity of  $G$  it is clear that  $G(t^+) = t^-$  and  $G(t^-) = t^+$ . By (5.6) we deduce that, if  $t_- < t_+$ , then

$$t^+ - t^- = G(t^-) - G(t^+) < t^+ - t^-,$$

which gives an absurd: hence,  $t^- = t^+$ , so  $t_j \rightarrow \bar{t}$  as we stated. Consider now any compact subset  $K$  of  $\Omega \setminus \{\tau = \bar{t}\}$ : there is then a small interval  $I \ni \bar{t}$  so that  $\tau(x) \notin I$  for all  $x \in K$ . There exists  $\bar{j}$ , depending on  $I$  thus on  $K$ , such that  $t_j \in I$  for all  $j \geq \bar{j}$ ; hence, for any  $x \in K$  and for any  $j \geq \bar{j}$ , one has  $\psi_j(x) = \bar{\psi}(x)$ .  $\square$

**Remark 5.2.** As the proof underlines, we have more than the uniform convergence of  $\psi_j$  to  $\bar{\psi}$  on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ : indeed, in such compact subsets one has that  $\psi_j \equiv \bar{\psi}$  for  $j \gg 1$ .

**Example 5.3** (Non-convergence). We show now that the *strict* 1-Lipschitzianity of  $G$  required in Proposition 5.1 is necessary. Indeed, let  $f$  and  $h_1, h_2$  be such that  $G(t) = 1 - t$  for  $t \in (-2, 2)$ : this is of course possible with a smooth choice of  $f$  and of  $h_1$  and  $h_2$ . Moreover, let  $\psi_0$  be given by (5.2): in this case it is clear that  $t_j = 0$  for all even  $j$  and  $t_j = 1$  for all odd  $j$ , so there is not convergence to  $\bar{t} = 1/2$ . Notice that in this case  $G$  is a Lipschitz function with Lipschitz constant equal to 1.

We can easily understand a practical meaning of the possible non-convergence that we have described: if a small town has two supermarkets, there is not a big difference for the people to drive to one or to the other one. However, it may happen that these supermarkets can have very long queues. Therefore, if in a certain day everybody decides to go to the first one, all the people will be unsatisfied with their choice, noticing that the other one has no queue. Consequently, in the next day everybody will go to the second supermarket, leading to a long queue in this one and no queue in the first one. Clearly this procedure will be repeated all the following days, and so there will be no convergence.

**Remark 5.4.** Let us underline what is the meaning of the request of 1-Lipschitzianity of  $G$ ; if everything is smooth, we can evaluate

$$G'(t) = -\left(h'_2(1 - m(t)) - h'_1(m(t))\right)m'(t);$$

hence, requiring that  $|G'|$  is small means that we want  $h'_i$  and/or  $m'$  small: the first fact means that the queue does not change dramatically for a small change in the quantity of people; and the second one means that  $f$  is not too big, so that a small region of the city (the region where the people can have the doubt whether going to one or the other place) does not contain too many citizens.

**5.2. Evolution with prudence.** As Example 5.3 shows, in the previous setting the convergence to the individual optimum does not always hold; however, this does not seem to be likely. In particular, in the case described in the second part of the example, it seems not so reasonable that people continue to change their idea every day in such a silly way. To model the “smarter” behaviour of the people, we can follow two ways. One way is to allow people to remember more than a single day, giving them a memory, and this is the content of Section 5.3. The other way, which is what we investigate now, is to make people more “prudent”.

To introduce the concept of prudence, we start by noticing that the easiness with which people change their habit is excessive: in other words, it is not very reasonable that for a generic  $x \in \Omega$  the function  $j \mapsto \psi_j(x)$  can change from 0 to 1 and vice-versa. The idea of “prudence” is that it can be more convenient to consider the choice of a citizen living at  $x$  at time  $j$  as a stochastic variable, hence using the approach with transport plans instead of transport maps; recalling (2.1), this means that the value of  $\psi_j(x)$  can be not only 0 or 1, but any number between 0 and 1. This approach is also common in game theory, and is related to the notion of mixed strategy, for which we refer to Section 7.4 of [5]. So,  $\psi_j(x)$  is the percentage of the citizens living at  $x$  which decide to go to  $x_2$  at the time  $j$ , while  $1 - \psi_j(x)$  is the percentage of those going to  $x_1$ . With this approach, one is not assuming that all the citizens living at the same place  $x$  are forced to take always exactly the same decisions. The evolution rule will be then a simple modification of (5.3). Recall the definition (5.1) of  $m_j$  which corresponds to the mass of people going to  $x_1$  at time  $j$ . Notice also that the definition (5.4) of  $t_{j+1}$  means that at time  $j + 1$ , for a citizen living at  $x$  it would seem at first glance more convenient to go to  $x_1$  (resp.  $x_2$ ) if  $\tau(x)$  is smaller (resp. greater) than  $t_{j+1}$ . We introduce then a parameter  $0 < \rho < 1$  which

we call *prudence parameter*, expressing the “resistance” that the average citizen feels against changing his habit. The evolution rule is then given as follows:

$$\psi_{j+1}(x) = \begin{cases} \rho\psi_j(x) & \text{if } \tau(x) < t_{j+1} \\ 1 - \rho(1 - \psi_j(x)) & \text{otherwise.} \end{cases} \quad (5.7)$$

Note that the extreme case  $\rho = 0$ , that is no prudence, corresponds to the case of the previous section; on the opposite, the extreme case  $\rho = 1$ , that is full prudence, means that nobody will ever change his habit, and so  $\psi_j = \psi_0$  for each  $j$ . We can then prove our result.

**Theorem 5.5.** *Assume that  $f$  is bounded and that  $h_1$  and  $h_2$  are non-decreasing and Lipschitz continuous. Then there exists  $0 < \bar{\rho} < 1$  so that, if  $\bar{\rho} < \rho < 1$ ,  $\psi_j$  converges to  $\bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

*Proof.* First of all, we introduce the quantities

$$S_j^2 = \int_{\{\tau(x) < \bar{t}\}} \psi_j(x) d\mu(x) \quad \text{and} \quad S_j^1 = \int_{\{\tau(x) > \bar{t}\}} (1 - \psi_j(x)) d\mu(x). \quad (5.8)$$

Observe that, at time  $j$ ,  $S_j^2$  is the amount of people which are going to  $x_2$ , but which should go to  $x_1$  according to the equilibrium  $\bar{\psi}$ ; analogously,  $S_j^1$  are the people going to  $x_1$  and which should go to  $x_2$ . Clearly, if we call  $\bar{m} = m(\bar{t})$ , it is  $m_j = \bar{m} + S_j^1 - S_j^2$ , where  $m_j$  is defined in (5.1). Hence, comparing (4.5) and (5.4) and recalling that  $h_1$  and  $h_2$  are Lipschitz, we deduce

$$|t_{j+1} - \bar{t}| \leq K |S_j^1 - S_j^2| \quad (5.9)$$

with a constant  $K$  depending only on  $h_1$  and  $h_2$ . Moreover, since  $h_1$  and  $h_2$  are non-decreasing and since  $G(\bar{t}) = \bar{t}$  by definition of equilibrium, we deduce that  $m_j \leq \bar{m}$  implies  $t_{j+1} \geq \bar{t}$  and symmetrically  $m_j \geq \bar{m}$  implies  $t_{j+1} \leq \bar{t}$ . This means that if at time  $j$  “few” people are going to  $x_1$ , at time  $j + 1$  “many” people will assume to be convenient to do so: however, thanks to the prudence, only a part of those people will indeed change their goal point.

Let us now take  $j \in \mathbb{N}$  and let us assume by symmetry that  $S_j^1 \geq S_j^2$ , hence  $m_j \geq m(\bar{t})$ ,  $t_{j+1} \leq \bar{t}$  and  $m(t_{j+1}) \leq m(\bar{t})$ . Keep in mind that, due to the prudence, it is not true in general that  $m_j = m(t_j)$ . By construction, we know

$$S_{j+1}^1 = \rho S_j^1. \quad (5.10)$$

On the other hand, concerning  $S_{j+1}^2$ , by definition (5.8) one has

$$\begin{aligned} S_{j+1}^2 &= \int_{\{\tau < \bar{t}\}} \psi_{j+1}(x) d\mu(x) = \int_{\{\tau < t_{j+1}\}} \psi_{j+1}(x) d\mu(x) + \int_{\{t_{j+1} < \tau < \bar{t}\}} \psi_{j+1}(x) d\mu(x) \\ &= \int_{\{\tau < t_{j+1}\}} \rho\psi_j(x) d\mu(x) + \int_{\{t_{j+1} < \tau < \bar{t}\}} (1 - \rho + \rho\psi_j(x)) d\mu(x) \\ &= \rho S_j^2 + (1 - \rho)(m(\bar{t}) - m(t_{j+1})) \leq \rho S_j^2 + (1 - \rho)D(S_j^1 - S_j^2). \end{aligned} \quad (5.11)$$

Here  $D = KL$ , where  $K$  is given by (5.9) and  $L$  is the Lipschitz constant of  $m$ : indeed, by immediate geometric arguments one understands that  $t \mapsto \mathbb{L}^d(\{x \in \Omega : \tau(x) \leq t\})$  is a Lipschitz map, so that  $m(t)$ , which is the integral over the sets  $\{x \in \Omega : \tau(x) \leq t\}$  of the bounded function  $f$ , is also Lipschitz.

Now, using (5.10) and (5.11) we have

$$S_{j+1}^1 - S_{j+1}^2 \geq (\rho - (1 - \rho)D)(S_j^1 - S_j^2).$$

We can finally deduce what follows: if the prudence is close enough to 1, namely  $\rho > 1 - \frac{1}{D+1}$ , the sign of  $S_m^1 - S_m^2$  remains positive for all  $m \geq j$ . In other words, if the population is prudent enough and at a certain moment too many people are going to the service  $x_1$ , then in the sequel the same service will always have more people than in the equilibrium, though the excess becomes smaller and smaller.

We can now observe that  $S_m^1$  and  $S_m^2$  go to 0 for  $m \rightarrow \infty$ . Indeed, by (5.10) we have that  $S_m^1 = \rho^{m-j} S_j^1$ , so  $S_m^1 \rightarrow 0$ ; moreover, being  $0 \leq S_m^2 \leq S_m^1$ , it is clearly also  $S_m^2 \rightarrow 0$ .

Using (5.9) we deduce that  $t_m \rightarrow \bar{t}$ , and finally by (5.7) this implies that  $\psi_m$  converges to  $\bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .  $\square$

**Example 5.6** (Non-convergence). We show that the claim of Theorem 5.5 is sharp, in the sense that a small prudence could be not enough to ensure convergence. As in the second part of Example 5.3, assume to be in a small city with two supermarkets having long queues, and take the prudence  $\rho = 1/3$ . Assume also that  $\psi_0(x) = 1/4$  for  $f$ -a.e.  $x$ : then, 25% of people are going to the first supermarket and 75% to the second one. Since the city is small, the queue convinces *all the citizens* that the first supermarket could be better; hence it will be  $\psi_1(x) = 3/4$  for  $f$ -a.e.  $x$ , so that at the day 1 one has 75% of people going to the first supermarket and 25% to the second one. Clearly, this procedure will be periodic and there will be no convergence to the equilibrium.

We conclude by noticing that this model with a “fixed prudence” may seem not so convincing, since it appears more likely that the prudence is increasing with the time: this increment can be also thought as a sign of “laziness”, or of habit. We describe now how a model with an increasing prudence (that is,  $\rho = \rho_j$  in the definition (5.7)) gives rise to convergence to the equilibrium.

**Theorem 5.7** (Increasing prudence). *Assume that  $f$  is bounded and that  $h_1$  and  $h_2$  are non-decreasing and Lipschitz continuous. Consider a sequence  $(\rho_j)_{j \geq 1}$  such that*

$$\rho_j \rightarrow 1 \quad \text{and} \quad \sum_{j=1}^{\infty} (1 - \rho_j) = +\infty.$$

*Consider an evolution with increasing prudence, i.e. consider (5.7) with  $\rho = \rho_j$ . Then there is convergence of  $\psi_j$  to  $\bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

*Proof.* Since  $\rho_j \rightarrow 1$ , there is  $\bar{j} \in \mathbb{N}$  so that for all  $j \geq \bar{j}$  one has  $\rho_j \geq \bar{\rho}$ , being  $\bar{\rho}$  as in Theorem 5.5. From the proof of that theorem, we know that, assuming by symmetry that  $S_j^1 \geq S_j^2$ , for all  $j \geq \bar{j}$  it is still  $S_j^1 \geq S_j^2$ . Notice also that a trivial calculation ensures

$$\prod_{j=\bar{j}}^{\infty} \rho_j = \exp \left( \ln \left( \prod_{j=\bar{j}}^{\infty} \rho_j \right) \right) = \exp \left( \sum_{j=\bar{j}}^{\infty} \ln \rho_j \right) \leq \exp \left( \sum_{j=\bar{j}}^{\infty} -(1 - \rho_j) \right) = 0. \quad (5.12)$$

Keeping in mind (5.10) and the fact that for all  $j \geq \bar{j}$  it is  $S_j^2 \leq S_j^1$ , we have that  $S_j^1 = S_j^1 \prod_{l=j}^{j-1} \rho_l \rightarrow 0$  and so also  $S_j^2 \rightarrow 0$ ; by (5.9), which does not depend on  $\rho_j$ , we deduce that  $t_j \rightarrow \bar{t}$ . Finally, the required convergence of  $\psi_j$  to  $\bar{\psi}$  is shown exactly as in Theorem 5.5 recalling (5.12).  $\square$

We now briefly discuss the assumptions of the above theorem: as already said, the meaning of  $\rho_j \rightarrow 1$  is that the people become more and more prudent, or more and more lazy. The fact that  $\sum_j (1 - \rho_j)$  must diverge tells us that if the prudence goes to 1 too fast, we can have

convergence to a situation different from the equilibrium: for instance, if  $\rho_j = 1$  for all  $j$ , as already pointed out we have the full prudence, so that the people do never change their habit, and so  $\psi_j = \psi_0$  for all  $j$ , whatever  $\psi_0$  is. Finally, we have this last remark.

**Remark 5.8.** Notice that the constant  $\bar{\rho}$  of Theorem 5.5 depends on the data of the problem. Since it is likely that the people only know how long the queues are everyday, but not the functions  $f$ ,  $h_1$  and  $h_2$ , a “smart” population does not have enough information to decide a fixed prudence  $\rho$  which ensures the convergence to the equilibrium. However, thanks to Theorem 5.7, the population can decide a strategy which surely leads to the convergence, that is an increasing prudence such as, for instance,  $\rho_j = 1 - 1/j$ .

**5.3. Evolution with memory.** We describe now the second possibility to overcome the unsatisfactory non-convergence behaviour of the simple evolution modeled in Section 5.1. As anticipated, the idea is to allow people to have a memory, so that every day they can decide what to do keeping in mind more than just the queue of the day before. We go back to the deterministic model, in which the position where a citizen lives completely determines his behaviour. Recall that the function  $G$  given by (5.5) represents the difference between the queues at  $x_2$  and at  $x_1$  if a quantity  $m(t)$  of people is going to  $x_1$ .

In the model of Section 5.1, each citizen assumes that at the day  $j + 1$  the queue would have been the same as in the day  $j$ . Here, we give to the people a “memory”: in other words, we take  $\kappa \in \mathbb{N}$ ,  $\kappa > 1$ , and in this model each citizen guesses that the queue will be approximatively the average of the queues of the preceding  $\kappa$  days. Therefore, the citizen assumes that the difference between the queues at  $x_2$  and at  $x_1$  at the day  $j + 1$  will be

$$Q = \frac{\sum_{m=j-\kappa+1}^j G(t_m)}{\kappa}.$$

As a consequence, the citizen living at  $x$  will decide to go to  $x_1$  at the day  $j + 1$  if  $\tau(x) < Q$ . We take  $\psi_1, \psi_2, \dots, \psi_\kappa$  as data of the problem, which means that for the first  $\kappa$  days we let the people decide freely how to move (as in Section 5.1, the choice of these  $\psi_j$ 's will not effect any of our results). For each  $n > \kappa$ , one has again

$$\psi_n(x) = \begin{cases} 0 & \text{if } \tau(x) < t_n \\ 1 & \text{otherwise,} \end{cases}$$

but this time the evolution rule is simply

$$t_{n+1} = \frac{\sum_{m=n-\kappa+1}^n G(t_m)}{\kappa}. \quad (5.13)$$

We give now the first result of this section, which reduces to Proposition 5.1 in the extreme case  $\kappa = 1$ .

**Theorem 5.9.** *Assume that  $h_1$  and  $h_2$  are continuous and non-decreasing. Assume that the function  $G$  given by (5.5) is Lipschitz with constant  $L < \kappa$ . Then  $t_n \rightarrow \bar{t}$  for  $n \rightarrow +\infty$ , hence also  $\psi_n \rightarrow \bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

To prove this theorem, we will need the following definition.

**Definition 5.10.** *Given  $n \in \mathbb{N}$  and  $j \in \{1, 2, \dots, \kappa\}$ , we say that the property  $P_n^j(\alpha)$  holds true when, given any set  $I \subseteq \{n, n-1, \dots, n-\kappa+1\}$  with  $\#I = j$  and  $n \in I$ , there holds*

$$\left| \sum_{i \in I} [G(t_i) - G(\bar{t})] \right| \leq \alpha. \quad (5.14)$$

The fundamental feature of the properties  $P_n^j(\alpha)$  is given by the lemma below.

**Lemma 5.11.** *Assume that  $G$  is  $L$ -Lipschitz with a constant  $L \leq \kappa$ . Then take  $n \in \mathbb{N}$  and  $\alpha > 0$ , and assume that  $P_m^j(\alpha)$  holds true for any  $j \in \{1, 2, \dots, \kappa\}$  and for any  $m \in \{n, n-1, \dots, n-\kappa+1\}$ . We have then what follows.*

**(A)** *For any  $j \in \{1, 2, \dots, \kappa\}$  and for any  $m > n$ , one has that  $P_m^j(\alpha)$  is true.*

**(B)** *For any  $j \in \{1, 2, \dots, \kappa\}$  and for any  $m > n + (j-1)\kappa$  one has that  $P_m^j(C_j\alpha)$  is true, where*

$$C_j = 1 - \left(1 - \frac{L}{\kappa}\right)^j. \quad (5.15)$$

Before proving the lemma, let us notice how the result of Theorem 5.9 easily follows.

*Proof of Theorem 5.9.* Let us set  $n = \kappa$ : there exists a positive  $\alpha$  such that all the properties  $P_m^j(\alpha)$  are valid for  $j$  and  $m$  in  $\{1, 2, \dots, \kappa\}$ , since there are a finite number of inequalities of the type (5.14) to be checked. By part **(A)** of Lemma 5.11 we know then that all the properties  $P_m^j(\alpha)$  hold true for any  $m \in \mathbb{N}$  and  $1 \leq j \leq \kappa$ . Moreover, since the constants  $C_j$  are increasing in  $j$  because  $L < \kappa$ , by part **(B)** we also have that the properties  $P_m^j(C_\kappa\alpha)$  are valid for any  $j$  and all  $m > \kappa + (\kappa-1)\kappa$ , that is

$$m \geq \kappa + \kappa(\kappa-1) + 1.$$

By an immediate iteration, we obtain the validity of the properties  $P_m^j(C_\kappa^h\alpha)$  for all  $j$  and all

$$m \geq \kappa + h(\kappa(\kappa-1) + 1).$$

Since  $C_\kappa < 1$ , the properties with  $j = 1$  are sufficient to deduce that  $G(t_i) - G(\bar{t}) \rightarrow 0$  for  $i \rightarrow \infty$ , and since  $G(t_i) = t_{i+1}$  while  $G(\bar{t}) = \bar{t}$  this ensures that  $t_i \rightarrow \bar{t}$  which is the first part of the thesis. The second part follows from the first one as already done in Proposition 5.1.  $\square$

Let us then prove the lemma.

*Proof of Lemma 5.11.* We start by proving **(A)**. Let us then fix  $\bar{m} > n$  and  $\bar{j}$ . We aim to show the validity of the property  $P_{\bar{m}}^{\bar{j}}(\alpha)$ . We can assume as an inductive hypothesis that  $P_m^j(\alpha)$  has already been established

- for all  $n < m < \bar{m}$  and all  $j$ ;
- for  $m = \bar{m}$  and all  $j < \bar{j}$ .

Let us now fix a set  $I$  according with Definition 5.10, and let us assume that  $t_{\bar{m}} \leq \bar{t}$  (so that  $G(t_{\bar{m}}) \geq G(\bar{t})$ ), which is clearly admissible by symmetry. Keeping in mind that  $G(\bar{t}) = \bar{t}$ , and then

$$t_{\bar{m}} - \bar{t} = \frac{\sum_{\{\bar{m}-\kappa \leq i < \bar{m}\}} [G(t_i) - G(\bar{t})]}{\kappa}, \quad (5.16)$$

we can first give the simple estimate

$$\sum_{i \in I} [G(t_i) - G(\bar{t})] = G(t_{\bar{m}}) - G(\bar{t}) + \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \geq \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \geq -\alpha. \quad (5.17)$$

Notice that the last inequality is empty true for  $\bar{j} = 1$ , even in the stronger form where 0 replaces  $-\alpha$ , while for  $\bar{j} > 1$  it is a consequence of the property  $P_m^{\bar{j}-1}(\alpha)$  with  $m = \max\{i \in I \setminus \{\bar{m}\}\}$ , which is already known to hold by assumption.

We must now show also the opposite inequality, that is,

$$\sum_{i \in I} [G(t_i) - G(\bar{t})] \leq \alpha. \quad (5.18)$$

To this aim notice that, calling  $J = \{\bar{m} - 1, \bar{m} - 2, \dots, \bar{m} - \kappa\} \setminus I$ , one can evaluate

$$\begin{aligned} t_{\bar{m}} - \bar{t} &= \frac{\left( \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \right) + \left( \sum_{i \in J} [G(t_i) - G(\bar{t})] \right)}{\kappa} \\ &\geq \frac{\left( \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \right) - \alpha}{\kappa}, \end{aligned} \quad (5.19)$$

using  $P_{m'}^j(\alpha)$  with  $m' = \max J < \bar{m}$  and  $j = \#J$ . Since in the same way one also has  $\sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \leq \alpha$ , the right term in last inequality is nonpositive: hence, by the  $L$ -Lipschitzianity of  $G$  and the fact that  $L \leq \kappa$  we obtain

$$G(t_{\bar{m}}) - G(\bar{t}) \leq L \frac{\alpha - \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})]}{\kappa} \leq \alpha - \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})], \quad (5.20)$$

which gives (5.18). Hence, we have shown **(A)** in its full generality.

Let us now attack **(B)**, which will be done performing the same estimates as above, but more carefully. Our proof will be obtained as an induction over  $\bar{j}$ , so we must start with the case  $\bar{j} = 1$ . In fact, this case is trivial: for any  $\bar{m} > n$ , recalling (5.16) we have

$$|t_{\bar{m}} - \bar{t}| = \left| \frac{\sum_{\{\bar{m}-\kappa \leq i < \bar{m}\}} [G(t_i) - G(\bar{t})]}{\kappa} \right| \leq \frac{\alpha}{\kappa}, \quad (5.21)$$

and hence by the  $L$ -Lipschitzianity of  $G$  one has

$$|G(t_{\bar{m}}) - G(\bar{t})| \leq \frac{L}{\kappa} \alpha;$$

this means that for any  $\bar{m} > n$  one has the validity of  $P_{\bar{m}}^1(C_1\alpha)$  with  $C_1 = \frac{L}{\kappa}$ , which coincides with (5.15).

We now take  $1 < \bar{j} \leq \kappa$ , suppose by induction that **(B)** has been proved for any  $j < \bar{j}$ , and try to obtain the thesis. We take then  $\bar{m} > n + (\bar{j} - 1)\kappa$ , a set  $I$  as in Definition 5.10 and, as before, to fix the ideas we start assuming without loss of generality that  $t_{\bar{m}} \leq \bar{t}$ . Since, as a consequence, we have  $G(t_{\bar{m}}) \geq G(\bar{t})$ , it is immediate to give the lower bound. Indeed, exactly as in (5.17), we can estimate

$$\begin{aligned} \sum_{i \in I} [G(t_i) - G(\bar{t})] &= G(t_{\bar{m}}) - G(\bar{t}) + \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \geq \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \\ &\geq -C_{\bar{j}-1}\alpha \geq -C_{\bar{j}}\alpha. \end{aligned} \quad (5.22)$$

To do so, we have used that  $C_{\bar{j}-1} \leq C_{\bar{j}}$ , which is clear by the definition (5.15), and we have also used the validity of  $P_{m'}^{\bar{j}-1}(C_{\bar{j}-1}\alpha)$  where  $m' = \max\{i \in I \setminus \{\bar{m}\}\}$ . And in turn,  $P_{m'}^{\bar{j}-1}(C_{\bar{j}-1}\alpha)$  is true by the inductive assumption and since

$$m' > \bar{m} - \kappa > n + (\bar{j} - 1)\kappa - \kappa = n + (\bar{j} - 2)\kappa.$$

To conclude the proof, we need then to show the upper bound, that is,

$$\sum_{i \in I} [G(t_i) - G(\bar{t})] \leq C_{\bar{j}}\alpha. \quad (5.23)$$

We start exactly as we did to show (5.18) in the proof of **(A)**; more precisely, we keep in mind the first inequality in (5.20), which has already been proved above, that is

$$G(t_{\bar{m}}) - G(\bar{t}) \leq L \frac{\alpha - \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})]}{\kappa}.$$

In order to use in a more careful way the  $L$ -Lipschitzianity of  $G$ , we rewrite the last inequality in the equivalent form

$$G(t_{\bar{m}}) - G(\bar{t}) \leq \frac{L}{\kappa} \alpha - \frac{L}{\kappa} \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})],$$

from which we obtain, using again some  $P_{m'}^{\bar{j}-1}(C_{\bar{j}-1}\alpha)$  of which we already know the validity by inductive assumption, that

$$\begin{aligned} \sum_{i \in I} [G(t_i) - G(\bar{t})] &\leq \frac{L}{\kappa} \alpha + \left(1 - \frac{L}{\kappa}\right) \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \\ &\leq \frac{L}{\kappa} \alpha + \left(1 - \frac{L}{\kappa}\right) \left| \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \right| \leq \frac{L}{\kappa} \alpha + \left(1 - \frac{L}{\kappa}\right) C_{\bar{j}-1} \alpha \quad (5.24) \\ &= \left[ \frac{L}{\kappa} + \left(1 - \frac{L}{\kappa}\right) \left(1 - \left(1 - \frac{L}{\kappa}\right)^{\bar{j}-1}\right) \right] \alpha = C_{\bar{j}} \alpha, \end{aligned}$$

hence (5.23) is established and we have completed the proof.  $\square$

As in the previous cases, we can give a counterexample to show that a small memory coefficient  $\kappa$  could be not enough to guarantee convergence.

**Example 5.12** (Non-convergence). This example is very similar to the ones in Examples 5.3 and 5.6: take a small city with the two supermarkets, and assume that the first supermarket can have a long queue, but the second one can have a much longer one. Take the memory coefficient  $\kappa = 2$ , and assume that in the days 0 and 1 everybody goes to the first supermarket. This gives, for the days  $j = 0$  and  $j = 1$ , an average queue which is high for the first supermarket and null for the second one: hence, at the day  $j = 2$  everybody will go to the second supermarket. The average queue for the days  $j = 1$  and  $j = 2$ , then, is high for the first shop but very high for the second, thus everybody will go back to the first one at the day  $j = 3$ . Again, the average queue for the days  $j = 2$  and  $j = 3$  will be high in the first supermarket and very high for the second one, so also at the day  $j = 4$  everybody will go to the first one. But then, at the day 5 all the people will move to the second one because the average queue for the days  $j = 3$  and  $j = 4$  is all at the first one. This procedure is clearly periodic with a period of 3 days and so the situation does not converge to the equilibrium. It is obvious how to modify the example to show that any memory coefficient  $\kappa$  can be not enough for a suitable choice of  $\Omega$ ,  $f$ ,  $h_1$  and  $h_2$ .

Let us give also a “precise” counterexample to show the exact role of the assumption  $L < \kappa$ .

**Example 5.13.** Let us take a situation so that  $G(t) = -2t$  for  $-3 \leq t \leq 3$ . Assume also that  $\kappa = 2$  and  $t_0 = t_1 = -1$ . Then it is immediate to deduce by (5.13) that it will be  $t_j = -1$  for all  $j \equiv 0 \pmod{3}$  or  $j \equiv 1 \pmod{3}$ , while  $t_j = 2$  for all  $j \equiv 2 \pmod{3}$ .

In the example above there is no convergence even though the Lipschitz constant  $L$  coincides with the memory coefficient  $\kappa$ , and it is immediate to modify the example to find a situation with no convergence with  $L = \kappa$  for any  $\kappa$ . This ensures that the assumption  $L < \kappa$  in Theorem 5.9

is almost sharp. Indeed, we can slightly strengthen the assumption of Theorem 5.9 covering also the situation in which  $L = \kappa$  but the Lipschitzianity is strict. We recall that a function  $\varphi : X \rightarrow Y$  is said *strictly Lipschitz with constant  $L$*  if for any  $x_1 \neq x_2$  in  $X$  one has

$$\|\varphi(x_1) - \varphi(x_2)\|_Y < L\|x_1 - x_2\|_X.$$

We have the following result.

**Theorem 5.14.** *Assume that  $h_1$  and  $h_2$  are continuous and non-decreasing and that the function  $G$  given by (5.5) is strictly Lipschitz with constant  $\kappa$ . Then  $t_n \rightarrow \bar{t}$  for  $n \rightarrow +\infty$ , hence also  $\psi_n \rightarrow \bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

*Proof.* We start by recalling that the result of Lemma 5.11 has been proved also in the case  $L = \kappa$ ; therefore, we can start as in the proof of Theorem 5.9 finding a constant  $\bar{\alpha}$  and knowing that all the properties  $P_m^j(\bar{\alpha})$  with  $1 \leq j \leq \kappa$  and  $m \in \mathbb{N}$  are valid. However, for  $L = \kappa$  all the constants  $C_j$  given in (5.15) are all equal to 1, so obtaining  $P_m^j(C_j^h \bar{\alpha})$  is of no help.

We now claim what follows: there exists a continuous and decreasing function  $\xi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  such that  $\xi(\alpha) < \alpha$  for all  $\alpha > 0$ , with the property that if all the  $P_m^j(\alpha)$  with  $1 \leq j \leq \kappa$  and  $m \geq n$  are valid, then all the  $P_m^j(\xi(\alpha))$  with  $1 \leq j \leq \kappa$  and  $m > n + \kappa(\kappa - 1)$  hold. Notice that this will immediately give the thesis, because for any  $h \in \mathbb{N}$  we will have, for  $m$  big enough, all the properties  $P_m^j(\xi^h(\alpha))$ , and by the continuity of  $\xi$  the sequence  $h \mapsto \xi^h(\bar{\alpha})$  must converge to 0.

To conclude the proof we take  $\alpha > 0$  and we suitably modify the proof of Lemma 5.11 for the case of  $L = \kappa$  but with the strict Lipschitzianity, so to show the existence of the desired constant  $\xi(\alpha) < \alpha$ . Our plan is to find constants  $\alpha_j < \alpha$  for  $1 \leq j \leq \kappa$  in such a way that  $P_m^j(\alpha_j)$  is true for  $m > \bar{m} + \kappa(j - 1)$ . The key idea is the following: since  $G$  is *strictly*  $\kappa$ -Lipschitz, for any  $\varepsilon \in ]0, \alpha[$  there exists  $\kappa_\varepsilon < \kappa$  so that

$$|G(t) - G(\bar{t})| \leq \kappa_\varepsilon |t - \bar{t}|, \quad \text{for all } t \in [\bar{t} - \alpha, \bar{t} - \varepsilon] \cup [\bar{t} + \varepsilon, \bar{t} + \alpha] : \quad (5.25)$$

this is trivial by compactness, but it will also be very useful for our purpose. We start arguing by induction. The case  $j = 1$  is very easy: indeed, we already noticed in (5.21) that for all  $\bar{m} > n$  one has  $|t_{\bar{m}} - \bar{t}| \leq \alpha/\kappa$ . Then we can fix an arbitrary  $\varepsilon < \alpha/\kappa$ . If  $|t_{\bar{m}} - \bar{t}| < \varepsilon$ , then by Lipschitzianity of  $G$  we have

$$|G(t_{\bar{m}}) - G(\bar{t})| \leq \kappa_\varepsilon;$$

on the other hand, if  $|t_{\bar{m}} - \bar{t}| \geq \varepsilon$ , by making use of (5.25) we know that

$$|G(t_{\bar{m}}) - G(\bar{t})| \leq \kappa_\varepsilon |t_{\bar{m}} - \bar{t}| \leq \frac{\kappa_\varepsilon}{\kappa} \alpha.$$

Hence we obtain the case  $j = 1$  with  $\alpha_1 = \min\{\kappa_\varepsilon \alpha / \kappa, \kappa_\varepsilon\} < \alpha$ .

Consider now the case  $1 < j \leq \kappa$ , and assume that the claim has been already proved up to  $j - 1$ . Let us begin as in Lemma 5.11: supposing that  $t_{\bar{m}} \leq \bar{t}$  by symmetry, as in (5.22) we have

$$\begin{aligned} \sum_{i \in I} [G(t_i) - G(\bar{t})] &= G(t_{\bar{m}}) - G(\bar{t}) + \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \\ &\geq \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \geq -\alpha_{j-1} \end{aligned} \quad (5.26)$$

and the estimate from below is done. Concerning the estimate from above, there are again two cases to consider: fix an arbitrary

$$\varepsilon < \frac{\alpha - \alpha_{j-1}}{\kappa}.$$

If  $|t_{\bar{m}} - \bar{t}| < \varepsilon$ , then

$$\sum_{i \in I} [G(t_i) - G(\bar{t})] \leq \kappa \varepsilon + \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \leq \kappa \varepsilon + \alpha_{j-1} < \alpha. \quad (5.27)$$

On the other hand, if  $|t_{\bar{m}} - \bar{t}| \geq \varepsilon$ , we can use (5.25) to derive, from (5.19) which says

$$t_{\bar{m}} - \bar{t} \geq \frac{\left( \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \right) - \alpha}{\kappa},$$

the estimate

$$G(t_{\bar{m}}) - G(\bar{t}) \leq \frac{\kappa \varepsilon}{\kappa} \left( \alpha - \left( \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \right) \right).$$

Therefore, we mimic (5.24) to find

$$\sum_{i \in I} [G(t_i) - G(\bar{t})] \leq \frac{\kappa \varepsilon}{\kappa} \alpha + \left( 1 - \frac{\kappa \varepsilon}{\kappa} \right) \sum_{i \in I \setminus \{\bar{m}\}} [G(t_i) - G(\bar{t})] \leq \frac{\kappa \varepsilon}{\kappa} \alpha + \left( 1 - \frac{\kappa \varepsilon}{\kappa} \right) \alpha_{j-1}. \quad (5.28)$$

Recalling (5.26), (5.27) and (5.28), we obtain the inductive step of the claim by setting

$$\alpha_j := \max \left\{ \alpha_{j-1}, \kappa \varepsilon + \alpha_{j-1}, \frac{\kappa \varepsilon}{\kappa} \alpha + \left( 1 - \frac{\kappa \varepsilon}{\kappa} \right) \alpha_{j-1} \right\} < \alpha.$$

We finally set  $\xi(\alpha) = \alpha_\kappa = \max_{j=1, \dots, \kappa} \alpha_j$ ; since the continuity of this map  $\xi$  is clear by the construction, the proof is obtained.  $\square$

We conclude the paper with a last result in the same philosophy of Theorem 5.7, in which we considered a prudence increasing in time: we consider what happens with an increasing memory. More precisely, we focus on the case of “global memory”, where the citizens remember *all* the previous days and then the definition (5.13) must be replaced by

$$t_{n+1} = \frac{\sum_{m=1}^n G(t_m)}{n}. \quad (5.29)$$

We can show that there is always convergence with global memory, and this leads exactly to the same considerations as in Remark 5.8.

**Theorem 5.15** (Global memory). *Assume that  $h_1$  and  $h_2$  are continuous and non-decreasing. Assume that the function  $G$  is Lipschitz, and that the evolution is given by the global memory (5.29). Then  $t_n \rightarrow \bar{t}$  and so  $\psi_n \rightarrow \bar{\psi}$  uniformly on the compact subsets of  $\Omega \setminus \{\tau = \bar{t}\}$ .*

*Proof.* We can rewrite (5.29) in a useful way as

$$t_{n+1} - \bar{t} = \left( 1 - \frac{1}{n} \right) (t_n - \bar{t}) + \frac{1}{n} (G(t_n) - G(\bar{t})).$$

Take then  $n \in \mathbb{N}$  and assume for simplicity that  $t_n \geq \bar{t}$ , so that  $G(t_n) \leq G(\bar{t})$ : we can estimate

$$\begin{aligned} t_{n+1} - \bar{t} &\leq \left( 1 - \frac{1}{n} \right) (t_n - \bar{t}), \\ t_{n+1} - \bar{t} &\geq \frac{1}{n} (G(t_n) - G(\bar{t})) \geq -\frac{L}{n} (t_n - \bar{t}); \end{aligned}$$

as a consequence, provided  $n$  is big enough, regardless of the sign of  $t_n - \bar{t}$  we have

$$|t_{n+1} - \bar{t}| \leq \left( 1 - \frac{1}{n} \right) |t_n - \bar{t}|.$$

Since  $\prod_{n=1}^{\infty} (1 - 1/n) = 0$ , this immediately leads to the thesis.  $\square$

**Remark 5.16.** One can of course consider a lot of other “schemes of memory”. For instance, one could have a global memory with different weights for older days, such as

$$t_{n+1} = \sum_{m=1}^n \frac{1}{2^m(1-2^{-n})} G(t_m),$$

where the term  $(1-2^{-n})$  is put to let the sum of the coefficients equal 1. It is not hard to notice that our method of proof can be easily adapted to a generality of those “schemes”: it is important that  $t_{n+1}$  is given by a mean of terms  $G(t_j)$  for some times  $j \leq n$  and that the associated coefficients are decreasing as the days become older; that is,

$$t_{n+1} = \sum_{m=1}^n C(n, m) G(t_j)$$

with  $0 \leq C(n, m) \leq 1$ ,  $\sum_{m=1}^n C(n, m) = 1$  and  $C(n, m) \leq C(n, m')$  for  $m < m'$ . In these cases one has the desired convergence provided  $G$  is  $L$ -Lipschitz with a suitable constant  $L$  as well as general convergence for a Lipschitz  $G$  for reasonable global schemes of memory.

## REFERENCES

- [1] L. Ambrosio, Lecture Notes on Optimal Transport Problems. Mathematical aspects of evolving interfaces, CIME Summer School in Madeira, vol. 1812, Springer, 2003.
- [2] L. Ambrosio, N. Fusco & D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems. Oxford University Press (2000).
- [3] L. Ambrosio, B. Kirchheim & A. Pratelli, Existence of optimal transport maps for crystalline norms. Duke Math. J. 125 (2004), no. 2, 207–241.
- [4] L. Ambrosio & A. Pratelli, Existence and stability results in the  $L^1$  theory of optimal transportation. Optimal transportation and applications (Martina Franca, 2001), 123–160, Lecture Notes in Math., 1813, Springer, Berlin, 2003.
- [5] J.-P. Aubin, Optima and equilibria. An introduction to nonlinear analysis. Translated from the French by Stephen Wilson. Second edition. Graduate Texts in Mathematics, 140. Springer-Verlag, Berlin, 1998. xviii+429 pp.
- [6] G. Bouchitté, C. Jimenez & M. Rajesh, Asymptotique d’un problème de positionnement optimal. (French) C. R. Math. Acad. Sci. Paris 335 (2002), no. 10, 853–858.
- [7] A. Brancolini, G. Buttazzo, F. Santambrogio & E. Stepanov, Long-term planning versus short-term planning in the asymptotical location problem. CVGMT preprint 2006, see <http://cvgmt.sns.it/papers/brabutsan06/>.
- [8] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. Comm. Pure Appl. Math. 44 (1991), no. 4, 375–417.
- [9] G. Buttazzo, A. Pratelli, S. Solimini & E. Stepanov, Optimal urban networks via mass transportation. Springer Lecture Notes in Math., to appear (2008), see <http://cvgmt.sns.it/papers/butprasol07/>.
- [10] G. Buttazzo, A. Pratelli & E. Stepanov, Optimal pricing policies for public transportation networks. SIAM J. Optim. 16 (2006), no. 3, 826–853.
- [11] G. Buttazzo & F. Santambrogio, A model for the optimal planning of an urban area. SIAM J. Math. Anal. 37 (2005), no. 2, 514–530.
- [12] L.A. Caffarelli, M. Feldman & R.J. McCann, Constructing optimal maps for Monge’s transport problem as a limit of strictly convex costs. J. Amer. Math. Soc. 15 (2002), no. 1, 1–26.
- [13] G. Carlier & F. Santambrogio, A variational model for urban planning with traffic congestion. ESAIM Control Optim. Calc. Var. 11 (2005), no. 4, 595–613.
- [14] G. Carlier & I. Ekeland, The structure of cities. J. Global Optim. 29 (2004), no. 4, 371–376.
- [15] G. Carlier & I. Ekeland, Equilibrium structure of a bidimensional asymmetric city. Nonlinear Anal. Real World Appl. 8 (2007), no. 3, 725–748.
- [16] L.C. Evans & W. Gangbo, Differential equations methods for the Monge-Kantorovich mass transfer problem. Mem. Amer. Math. Soc. 137 (1999), no. 653, viii+66 pp.
- [17] H. Federer, Geometric measure theory. Springer (1969).

- [18] A. Friedman, Differential games. Pure and Applied Mathematics, Vol. XXV. Wiley-Interscience, New York-London, 1971. xii+350 pp.
- [19] W. Gangbo & R. J. McCann, The geometry of optimal transportation. *Acta Math.* 177 (1996), no. 2, 113–161.
- [20] C. Jimenez, Optimisation de problèmes de transport. PhD Thesis, Univ. du Sud-Toulon Var, (2005)
- [21] L.V. Kantorovich, On the transfer of masses. *Dokl. Akad. Nauk. SSSR* 37 (1942), 227–229.
- [22] L.V. Kantorovich, On a problem of Monge. *Uspekhi Mat. Nauk.* 3 (1948), 225–226.
- [23] M. Knott & C.S. Smith, On the optimal mapping of distributions. *J. Optim. Theory Appl.* 43 (1984), no. 1, 39–49.
- [24] R.E. Lucas & Jr., E. Rossi-Hansberg, On the internal structure of cities. *Econometrica* 70 (2002), no. 4, 1445–1476.
- [25] G. Monge, Memoire sur la Theorie des Déblais et des Remblais. *Histoire de l’Acad. de Sciences de Paris*, 1781.
- [26] L. Rüschendorf, Optimal solutions of multivariate coupling problems. *Appl. Math. (Warsaw)* 23 (1995), no. 3, 325–338.
- [27] C. Villani, Topics in Optimal transportation. Graduate studies in Mathematics, 58, AMS (2003).

G.C.: DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI PARMA, VIALE G.P. USBERTI 53/A (CAMPUS), 43100 PARMA, ITALY

*E-mail address:* `gianluca.crippa@unipr.it`

C.J.: LABORATOIRE DE MATHÉMATIQUES DE BREST, 6, AVENUE VICTOR LE GORGEU, CS 93837, F-29238 BREST CEDEX 3, FRANCE

*E-mail address:* `chloe.jimenez@univ-brest.fr`

A.P.: DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI PAVIA, VIA FERRATA 1, 27100 PAVIA, ITALY

*E-mail address:* `aldo.pratelli@unipv.it`